SOLUTIONS TO THE MULTIVARIATE G-SAMPLE BEHRENS-FISHER PROBLEM BASED UPON GENERALIZATIONS OF THE BROWN-FORSYTHE $\underline{F}^*$ AND WILCOX $\underline{H}_m$ TESTS

By

WILLIAM THOMAS COOMBS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1992

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to the individuals that have assisted me in completing this study. First, I would like to thank Dr. James J. Algina, chairperson of my doctoral committee, for (a) suggesting the topic for my dissertation, (b) guiding me through difficult applied and theoretical barriers, (c) debugging computer errors, (d) providing editorial suggestions, and (e) fostering my professional and personal growth through encouragement, support, and friendship. Second, I am indebted and grateful to the other members of my committee, Dr. Linda M. Crocker, Dr. M. David Miller, and Dr. Ronald H. Randles, for patiently reading the manuscript, offering constructive suggestions, providing editorial assistance, and giving continuous support. Third, I must thank Dr. John M. Newell who as a fifth and unofficial member of my committee still attended committee meetings, read the manuscript, and vigilantly inquired as to the progress of the project. Finally, I would like to express my heartfelt thanks to my wife Laura and son Tommy. Space limitations prevent me from enumerating the many personal sacrifices, both large and small, required of my wife so that I was able to accomplish this task. Although I shall never be able to fully repay the debt I have incurred to both my

committee and family, let me begin simply and sincerely--thank you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

ABSTRACT vi

CHAPTERS

1 INTRODUCTION 1

The Problem 3
Purpose of the Study 4
Significance of the Study 4

2 REVIEW OF LITERATURE 9

The Independent Samples *t* Test 9
Alternatives to the Independent Samples *t* Test . . 10
ANOVA *F* Test 15
Alternatives to the ANOVA *F* Test 16
Hotelling's T^2 Test 26
Alternatives to the Hotelling's T^2 Test 30
MANOVA Criteria 32
Alternatives to the MANOVA Criteria 36

3 METHODOLOGY 39

Development of Test Statistics 39
Brown-Forsythe Generalizations 39
Scale of the measures of between and within group variability 43
Equality of expectation of the measures of between and within group dispersion 47
Wilcox Generalization 49
Invariance Property of the Test Statistics 50
Brown-Forsythe Generalizations 51
Wilcox Generalization 52
Design 54
Simulation Procedure 62
Summary 65

4 RESULTS AND DISCUSSION 67

Brown-Forsythe Generalizations 74
Johansen Test and Wilcox Generalization 90

5 CONCLUSIONS 96

General Observations 96
Suggestions to Future Researchers 97

APPENDIX ESTIMATED TYPE I ERROR RATES 100

REFERENCES 141

BIOGRAPHICAL SKETCH 148

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

SOLUTIONS TO THE MULTIVARIATE G-SAMPLE BEHRENS-FISHER PROBLEM BASED UPON GENERALIZATIONS OF THE BROWN-FORSYTHE $\underline{F}^*$ AND WILCOX $\underline{H}_m$ TESTS

By

William Thomas Coombs

August, 1992

Chairperson: James J. Algina
Major Department: Foundations of Education

The Brown-Forsythe $\underline{F}^*$ and Wilcox $\underline{H}_m$ tests are generalized to form multivariate alternatives to MANOVA for use in situations where dispersion matrices are heteroscedastic. Four generalizations of the Brown-Forsythe $\underline{F}^*$ test are included.

Type I error rates for the Johansen test and the five new generalizations were estimated using simulated data for a variety of conditions. The design of the experiment was a 2^8 factorial. The factors were (a) type of distribution, (b) number of dependent variables, (c) number of groups, (d) ratio of total sample size to number of dependent variables, (e) form of the sample size ratio, (f) degree of the sample size ratio, (g) degree of heteroscedasticity, and (h) relationship of sample size to dispersion matrices. Only conditions in

which dispersion matrices were heterogeneous were included.

In controlling Type I error rates, the four generalizations of the Brown-Forsythe $\underline{F}^*$ test greatly outperform both the Johansen test and the generalization of the Wilcox $\underline{H}_m$ test.

CHAPTER 1
INTRODUCTION

Comparing two population means by using data from independent samples is one of the most fundamental problems in statistical hypothesis testing. One solution to this problem, the independent samples *t* test, is based on the assumption that the samples are drawn from populations with equal variances. According to Yao (1965), Behrens (1929) was the first to solve testing H_o: $\mu_1 = \mu_2$ without making the assumption of equal population variances. Fisher (1935, 1939) showed that Behrens solution could be derived from Fisher's theory of statistical inference called fiducial probability. Others (Aspin, 1948; Welch, 1938, 1947) have proposed solutions to the two-sample Behrens-Fisher problem as well.

The independent samples *t* test has been generalized to the analysis of variance (ANOVA) *F* test, a test of the equality of G population means. This procedure assumes homoscedasticity, that is, $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_G^2$. Several authors have proposed procedures to test H_o: $\mu_1 = \mu_2 = \ldots = \mu_G$ without assuming equal population variances. Welch (1951) extended his 1938 work and arrived at an approximate degrees of freedom (APDF) solution. Brown and Forsythe (1974), James

(1951), and Wilcox (1988, 1989) have proposed other solutions to the G-sample Behrens-Fisher problem.

Hotelling (1931) generalized the independent samples t test to a test of the equality of two population mean vectors. This procedure makes the assumption of equal population dispersion (variance-covariance) matrices, that is, $\Sigma_1 = \Sigma_2$. Several authors have proposed procedures to test H_o: $\mu_1 = \mu_2$ without assuming equal population dispersion matrices. James (1954) generalized his 1951 work and arrived at a series solution. Anderson (1958), Bennet (1951), Ito (1969), Nel & van der Merwe (1986), Scheffe (1943), and Yao (1965) have proposed additional solutions to the multivariate two-sample Behrens-Fisher problem.

Bartlett (1939), Hotelling (1951), Lawley (1938), Pillai (1955), Roy (1945), and Wilks (1932) have proposed multivariate generalizations of the ANOVA F test, creating the four basic multivariate analysis of variance (MANOVA) procedures for testing H_o: $\mu_1 = \mu_2 = \ldots = \mu_G$. These procedures make the assumption of equal population dispersion matrices. James (1954) and Johansen (1980) proposed procedures to test the equality of G mean vectors without making the assumption of homoscedasticity, that is, $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_G$. James extended James's (1951) univariate procedures to produce first-order and second-order series solutions. Johansen generalized the Welch (1951) procedure to form an APDF solution to this problem.

The Problem

To date, neither the Brown-Forsythe (1974) nor the Wilcox (1989) procedure has been extended to the multivariate setting. To test H_o: $\mu_1 = \mu_2 = \ldots = \mu_G$ Brown and Forsythe (1974) proposed the statistic

$$F^* = \frac{\sum_{i=1}^{G} n_i(\bar{x}_{i.} - \bar{x}_{..})^2}{\sum_{i=1}^{G} (1 - \frac{n_i}{N}) s_i^2}$$

where n_i denotes the number of observations in the ith group, $\bar{x}_{i.}$ the mean for the ith group, $\bar{x}_{..}$ the grand mean, s_i^2 the variance of the ith group, N the total number of observations, and G the number of groups. The statistic $\underline{F}^*$ is approximately distributed as $\underline{F}$ with G-1 and f degrees of freedom, where

$$f = \frac{[\sum_{i=1}^{G} (1 - \frac{n_i}{N}) s_i^2]^2}{\sum_{i=1}^{G} \frac{[(1 - \frac{n_i}{N}) s_i^2]^2}{n_i - 1}} .$$

The degrees of freedom, f, were determined by using a procedure due to Satterthwaite (1941).

To test H_o : $\mu_1 = \mu_2 = \ldots = \mu_G$ Wilcox (1989) proposed the statistic

$$H_m = \sum_{i=1}^{G} w_i(\tilde{x}_i - \tilde{x})^2 ,$$

where

$$w_i = [\frac{s_i^2}{n_i}]^{-1}$$

$$\tilde{X}_i = \frac{2x_{in_i}}{n_i(n_i + 1)} + \frac{n_i - 1}{n_i(n_i + 1)}\bar{X}_i.$$

and

$$\tilde{X} = \sum_{i=1}^{G} \frac{w_i \tilde{X}_i}{\sum_{i=1}^{G} w_i} \quad .$$

In the equation for $\tilde{X}_i$, x_{in_i} denotes the last observation in the ith sample. The statistic $\underline{H}_m$ is approximately distributed as chi-square with G-1 degrees of freedom.

Purpose of the Study

The purpose of this study is to extend the univariate procedures proposed by Brown and Forsythe (1974) and Wilcox (1989) to test $H_0 : \mu_1 = \mu_2 = \ldots = \mu_G$ and to compare Type I error rates of the proposed multivariate generalizations to the error rates of Johansen's (1980) test under varying distributions, numbers of dependent (criterion) variables, numbers of groups, forms of the sample size ratio, degrees of the sample size ratio, ratios of total sample size to number of dependent variables, degrees of heteroscedasticity, and relationships of sample size to dispersion matrices.

Significance of the Study

The application of multivariate analysis of variance in education and the behavioral sciences has increased dramatically, and it appears that it will be used frequently

in the future for data analysis (Bray & Maxwell, 1985, p.7). Stevens suggested three reasons why multivariate analysis is prominent:

> 1. Any worthwhile treatment will affect the subject in more than one way, hence the problem for the investigator is to determine in which specific ways the subjects will be affected and then find sensitive measurement techniques for those variables.
>
> 2. Through the use of multiple criterion measures we can obtain a more complete and detailed description of the phenomenon under investigation.
>
> 3. Treatments can be expensive to implement, while the cost of obtaining data on several dependent variables is relatively small and maximizes information gain. (1986, p. 2)

Hotelling's $\underline{T}^2$ is sensitive to violations of homoscedasticity, particularly when sample sizes are unequal (Algina & Oshima, 1990; Algina, Oshima, & Tang, 1991; Hakstian, Roed, & Lind, 1979; Holloway & Dunn, 1967; Hopkins & Clay, 1963; Ito & Schull, 1964). Yao's (1965), James's (1954) first- and second-order, and Johansen's (1980) tests are alternatives to Hotelling's $\underline{T}^2$ that have no underlying assumption of homoscedasticity. In controlling Type I error rates under heteroscedasticity, Yao's test is superior to James's first-order test (Algina & Tang, 1988; Yao, 1965). Algina, Oshima, and Tang (1991) studied Type I error rates of the four procedures when applied to data sampled from multivariate distributions composed of p independent univariate distributions. When (a) sample sizes are unequal and (b) dispersion matrices are unequal, the four procedures

can be seriously nonrobust with extremely skewed distributions such as the exponential and lognormal, but are fairly robust with moderately skewed distributions such as the beta(5,1.5). They also appear to be robust with non-normal symmetric distributions such as the uniform, $\underline{t}$, and Laplace. The performance of Yao's test, James's second-order test, and Johansen's test was slightly superior to the performance of James's first-order test (Algina, Oshima, & Tang, 1991).

MANOVA criteria are relatively robust to non-normality (Olson, 1974, 1976) but are sensitive to violations of homoscedasticity (Korin, 1972; Olson, 1974, 1979; Pillai & Sudjana, 1975; Stevens, 1979). The Pillai-Bartlett trace criterion is the most robust of the four basic MANOVA criteria for protection against non-normality and heteroscedasticity of dispersion matrices (Olson, 1974, 1976, 1979). Alternatives to MANOVA criteria that are not based on the homoscedasticity assumption include James's first- and second-order tests, and Johansen's test. When (a) sample sizes are unequal, (b) dispersion matrices are unequal, and (c) data are sampled from multivariate normal distributions, Johansen's test and James's second-order test outperform the Pillai-Bartlett trace criterion and James's first-order test (Tang, 1989).

In the univariate case, the Brown-Forsythe $\underline{F}^*$ test and Wilcox $\underline{H}_m$ test do not require the equality of population variances for the G groups. Hence, the Brown-Forsythe and Wilcox tests are more general procedures than the ANOVA $\underline{F}$

test. This suggests that generalizations of the Brown-Forsythe procedure and the Wilcox procedure might have advantages over the commonly used MANOVA procedure in cases of heteroscedasticity.

Brown and Forsythe (1974) used Monte Carlo techniques to examine the ANOVA $\underline{F}$ test, Brown-Forsythe $\underline{F}^*$ test, Welch APDF test, and James first-order procedure. The critical value proposed by Welch is a better approximation for small sample size than that proposed by James. Under (a) normality and (b) inequality of variances both Welch's test and the $\underline{F}^*$ test tend to have actual Type I error rates (τ) near nominal error rates (α) in a wide variety of conditions. However, there are conditions in which each fails to control τ. In terms of power, the choice between Welch (the specialization of Johansen's test and, in the case of two groups, of Yao's test) and $\underline{F}^*$ depends upon the magnitude of the means and their standard errors. The Welch test is preferred to the $\underline{F}^*$ test if extreme means coincide with small variances. When the extreme means coincide with large variances, the power of the $\underline{F}^*$ test is greater than that of the Welch test.

A limited simulation by Clinch and Keselman (1982) indicated that under conditions of heteroscedasticity, the Brown-Forsythe test is less sensitive to non-normality than is Welch's test. In fact, Clinch and Keselman concluded the user should uniformly adopt the $\underline{F}^*$ test over the Welch test. More recently Oshima and Algina (in press) reported that, with non-

normal data, in some conditions the $\underline{F}^*$ test has better control over τ than does James's second-order test, Welch's test, or Wilcox's $\underline{H}_m$ test. In other conditions the $\underline{F}^*$ test has substantially worse control. Oshima and Algina concluded that James's second-order test should be used with symmetric distributions and Wilcox's $\underline{H}_m$ test should be used with moderately asymmetric distributions. With markedly asymmetric distributions none of the tests had good control of τ. Extensive simulations (Wilcox, 1988) indicated that under normality the Wilcox $\underline{H}$ procedure always gave the experimenter more control over Type I error rates than did the $\underline{F}^*$ or Welch test and has error rates similar to James's second-order method, regardless of the degree of heteroscedasticity. Wilcox (1989) proposed $\underline{H}_m$, an improvement to the Wilcox (1988) $\underline{H}$ method; the improved test is much easier to use than James's second-order method. Wilcox (1990) indicated that the $\underline{H}_m$ test is more robust to non-normality than is the Welch test. Because the Johansen (1980) procedure is the extension of the Welch test, the results reported by Clinch and Keselman and by Wilcox suggest generalizations of the Brown-Forsythe procedure and the Wilcox procedure might have advantages over the Johansen procedure in some cases of heteroscedasticity and/or skewness. Thus, the construction and comparison of new procedures which may be competitive or even superior under some conditions than the established standard is merited.

CHAPTER 2
REVIEW OF LITERATURE

Independent Samples t Test

The independent samples _t_ is used to test the hypothesis of the equality of two population means when independent random samples are drawn from two populations which are normally distributed and have equal population variances. The test statistic

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$s_p^2 = \frac{(n_1-1)\,s_1^2 + (n_2-1)\,s_2^2}{n_1+n_2-2}$$

has a _t_ distribution with n_1+n_2-2 degrees of freedom.

The degree of robustness of the independent samples _t_ test to violations of the assumption of homoscedasticity has been well documented (Boneau, 1960; Glass, Peckham, & Sanders, 1972; Holloway & Dunn, 1967; Hsu, 1938; Scheffe, 1959). In cases where there are unequal population variances, the relationship between the actual Type I error rate (τ) and the nominal Type I error rate (α) is influenced by the sample size. When sample sizes are equal ($n_1=n_2$) and sufficiently

large, τ and α are near one another. In fact, Scheffe (1959, p.339) has shown for equal-sized samples t is asymptotically standard normal, even though the two populations are non-normal or have unequal variances. However, Ramsey (1980) found there are boundary conditions where t is no longer robust to violations of homoscedasticity even with equal-sized samples selected from normal populations. Results from numerous studies (Boneau, 1960; Hsu, 1938; Pratt, 1964; Scheffe, 1959) have shown that when the sample sizes are unequal and the larger sample is selected from the population with larger variance (known as the positive condition), the t test is conservative (that is, $\tau < \alpha$). Conversely, when the larger sample is selected from the population with smaller variance (known as the negative condition), the t test is liberal (that is, $\tau > \alpha$).

Alternatives to the Independent Samples t Test

According to Yao (1965), Behrens (1929) was the first to propose a solution to the problem of testing the equality of two population means without assuming equal population variances. This problem has come to be known as the Behrens-Fisher problem. Fisher (1935,1939) noted that Behrens's solution could be derived using Fisher's concept of fiducial distributions.

A number of other tests have been developed to test the hypothesis H_o: $\mu_1 = \mu_2$ in situations in which $\sigma_1^2 \neq \sigma_2^2$. Welch (1947) reported several tests in which the test statistic is

$$t_v = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad .$$

The critical value is different for the various tests. There are two types of critical values: (a) approximate degrees of freedom (APDF), and (b) series.

The APDF critical value (Welch, 1938) is a fractile of Student's t distribution with

$$f = \frac{\left[\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right]^2}{\frac{\left(\frac{\sigma_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{\sigma_2^2}{n_2}\right)^2}{n_2 - 1}}$$

degrees of freedom. In practice, the estimator of f is obtained by replacing parameters by statistics, that is, s_i^2 replaces σ_i^2 (i=1,2). In the literature the test using this estimator for f is referred to as the Welch test.

Welch (1947) expressed the series critical value for $\underline{t}_v$ as a function of s_1^2, s_2^2, and α, and developed a series critical value in powers of $(n_i - 1)^{-1}$. The first three terms in the series critical value are shown in Table 1. The zero-order term is simply a fractile of the standard normal distribution (z); using the zero-order term as the critical value is appropriate with large samples. The first-order critical value is the sum of the zero- and first-order terms,

Table 1

Critical Value Terms for Welch's (1947) Zero-, First-, and Second-Order Series Solutions

Power of $(n_i - 1)^{-1}$	Term
Zero	z
One	$z\left[\frac{1+z^2}{4}\frac{\sum_{i=1}^{2}\frac{\left(\frac{s_i^2}{n_i}\right)^2}{n_i-1}}{\left[\sum_{i=1}^{2}\frac{s_i^2}{n_i}\right]^2}\right]^{\frac{1}{2}}$
Two	$z\left[-\frac{1+z^2}{2}\frac{\sum_{i=1}^{2}\left(\frac{s_i^2}{n_i(n_i-1)}\right)^2}{\left[\sum_{i=1}^{2}\frac{s_i^2}{n_i}\right]^2}\right.$
	$+\frac{3+5z^2+z^4}{3}\frac{\sum_{i=1}^{2}\left[\frac{\left(\frac{s_i^2}{n_i}\right)^3}{n_i-1}\right]^2}{\left[\sum_{i=1}^{2}\frac{s_i^2}{n_i}\right]^3}$
	$\left.-\frac{15+32z^2+9z^4}{32}\frac{\sum_{i=1}^{2}\left[\frac{\left(\frac{s_i^2}{n_i}\right)^2}{n_i-1}\right]^2}{\left[\sum_{i=1}^{2}\frac{s_i^2}{n_i}\right]^4}\right]^{\frac{1}{2}}$

Note. Z denotes a fractile of the normal distribution.

whereas the second-order critical value is the sum of all three terms. As the sample sizes decline, there is a greater need for the more complicated critical values. James (1951) and James (1954) generalized the Welch series solutions to the G-sample case and multivariate cases respectively. Consequently, tests using the series solution are referred to as James's first-order and second-order tests. The zero-order test is often referred to as the asymptotic test. Aspin (1948) reported the third- and fourth-order terms, and investigated, for equal-sized samples, variation in the first- through fourth-order critical values.

Wilcox (1989) proposed a modification to the asymptotic test. The Wilcox statistic

$$\frac{\tilde{x}_1-\tilde{x}_2}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}\ ,$$

where

$$\tilde{x}_i = \frac{2x_{in_i}}{n_i(n_i+1)} + \frac{n_i-1}{n_i(n_i+1)}\bar{x}_i\ ,$$

is asymptotically distributed as a standard normal distribution. Here $\tilde{x}_i$ $(i=1,2)$ are biased estimators of the population means which result in improved empirical Type I error rates (Wilcox, 1989).

The literature suggests the following conclusions in the two-sample case regarding the control of Type I error rates under normality and heteroscedasticity for the independent samples <u>t</u> test, Welch APDF test, James first- and second-order

series tests, Brown-Forsythe test, and Wilcox $\underline{H}_m$ test: (a) the performance of the Welch test and Brown-Forsythe test is superior to the $\underline{t}$ test; (b) the Wilcox test and James second-order test are superior to the Welch APDF test; and (c) for most applications in education and the social sciences where data are sampled from normal distributions under heteroscedasticity, the Welch APDF test is adequate. Scheffe (1970) examined six different tests including the Welch APDF test from the standpoint of the Neyman-Pearson school of thought. Scheffe concluded the Welch test, which requires only the easily accessible $\underline{t}$-table, is a satisfactory practical solution to the Behrens-Fisher problem. Wang (1971) examined the Behrens-Fisher test , Welch APDF test, and Welch-Aspin series test (Aspin, 1948; Welch, 1947). Wang found the Welch APDF test to be superior to the Behrens-Fisher test when combining over all the experimental conditions considered. Wang found $|\tau-\alpha|$ was smaller for the Welch-Aspin series test than for the Welch APDF test. Wang noted, however, that the Welch-Aspin series critical values were limited to a select set of sample sizes and nominal Type I error rates. Wang concluded, in practice, one can just use the usual $\underline{t}$-table to carry out the Welch APDF test without much loss of accuracy. However, the Welch APDF test becomes conservative with very long-tailed symmetric distributions (Yuen, 1974). Wilcox (1990) investigated the effects of non-normality and heteroscedasticity on the Wilcox and Welch APDF tests. The

Wilcox test tended to outperform the Welch test. Moreover, over all conditions, the range of τ was (.032, .065) for α=.05, indicating the Wilcox test may have appropriate Type I error rates under heteroscedasticity and non-normality.

In summary, the independent samples t test is generally acceptable in terms of controlling Type I error rates provided there are sufficiently large equal-sized samples, even when the assumption of homoscedasticity is violated. For unequal-sized samples, however, an alternative that does not assume equal population variances such as the Wilcox test or the James second-order series test is preferable.

ANOVA F Test

The ANOVA F is used to test the hypothesis of the equality of G population means when independent random samples are drawn from populations which are normally distributed and have equal population variances. The test statistic

$$F = \frac{\sum_{i=1}^{G} n_i(\overline{x}_{i.}-\overline{x}_{..})^2/(G-1)}{\sum_{i=1}^{G} (n_i-1)\, s_i^2/(N-G)}$$

has an F distribution with G-1 and N-G degrees of freedom.

Numerous studies have shown that the ANOVA F test is not robust to violations of the assumption of homoscedasticity (Clinch & Keselman, 1982; Brown & Forsythe, 1974; Kohr & Games, 1974; Rogan & Keselman, 1977; Wilcox, 1988). The behavior of ANOVA F parallels that of the independent samples

t test with one exception. Whereas the independent samples t is generally robust when large sample sizes are equal, the ANOVA F may not maintain adequate control of Type I error rates even with equal-sized samples if the degree of heteroscedasticity is large (Rogan & Keselman, 1977; Tomarken & Serlin, 1986). In the positive condition the F test is conservative and in the negative condition the F test is liberal (Box, 1954; Clinch & Keselman, 1982; Brown & Forsythe, 1974; Horsnell, 1953; Rogan & Keselman, 1972; Wilcox, 1988).

Alternatives to the ANOVA F Test

A number of tests have been developed to test the hypothesis $H_o: \mu_1 = \mu_2 = \ldots = \mu_G$ in situations in which $\sigma_i^2 \neq \sigma_j^2$ (for at least one pair of i and j). Welch (1951) generalized the Welch (1938) APDF solution and proposed the statistic

$$F_v = \frac{\sum_{i=1}^{G} w_i(\overline{x}_i - \overline{x})^2/(G-1)}{1+\frac{2(G-2)}{G^2-1}\sum_{i=1}^{G}\frac{1}{f_i}(1-\frac{w_i}{w})^2}$$

where

$$w_i = [\frac{s_i^2}{n_i}]^{-1} \quad i=1,\ldots,G$$

$$w = \sum_{i=1}^{G} w_i$$

$$\overline{x}_i = \frac{1}{n_i}\sum_{i=1}^{n_i} x_{ij} \quad i=1,\ldots,G$$

$$\overline{x} = \sum_{i=1}^{G} \frac{w_i \overline{x}_i}{w}$$

and

$$f_i = n_i - 1 \quad i=1,\ldots,G \quad .$$

The statistic $\underline{F}_v$ is approximately distributed as $\underline{F}$ with G-1 and

$$\left[\frac{3}{G^2-1} \sum_{i=1}^{G} \frac{1}{f_i} \left(1 - \frac{w_i}{w}\right)^2 \right]^{-1}$$

degrees of freedom.

James (1951) generalized the Welch (1947) series solutions, proposing the test statistic

$$J = \sum_{i=1}^{G} w_i (\overline{x}_i - \overline{x})^2$$

where

$$w_i = \left[\frac{s_i^2}{n_i} \right]^{-1} \quad i=1,\ldots,G$$

$$\overline{x}_i = \frac{1}{n_i} \sum_{i=1}^{n_i} x_{ij} \quad i=1,\ldots,G$$

$$\overline{x} = \sum_{i=1}^{G} \frac{w_i \overline{x}_i}{w}$$

and

$$w = \sum_{i=1}^{G} w_i \quad .$$

In the asymptotic test the critical value of the statistic $\underline{J}$ is a fractile of a chi-square distribution with G-1 degrees of

freedom. If the sample sizes are not sufficiently large, however, the distribution of the test statistic may not be accurately approximated by a chi-square distribution with G-1 degrees of freedom. James (1951) derived a series expression which is a function of the sample variances such that

$$P\ [\sum_{i=1}^{G} w_i(\overline{x}_i - \tilde{x})^2 \geq 2h(s_i^2)\ ] \ = \ \alpha \quad .$$

James found approximations to $2h(s_i^2)$ of orders $1/f_i$ and $1/f_i^2$ ($f_i = n_i - 1$). In the first-order test, James found to order $1/f_i$ the critical value is

$$2h(s_i^2) \ = \ \chi^2_{G-1;\alpha}\ [\ 1 + \frac{3\ \chi^2_{G-1;\alpha} + G + 1}{2(G^2-1)} \sum_{i=1}^{G} \frac{1}{f_i}(1 - \frac{w_i}{w})^2\ ] \quad .$$

The null hypothesis is rejected in favor of the alternative hypothesis if $\underline{J} \geq 2h(s_i^2)$. James also provided a second-order solution which approximates $2h(s_i^2)$ to order $1/f_i^2$. James noted that this second-order test is very computationally intensive.

Brown and Forsythe (1974) proposed the test statistic

$$F^* \ = \ \frac{\sum_{i=1}^{G} n_i(\overline{x}_{i.} - \overline{x}_{..})^2}{\sum_{i=1}^{G} (1 - \frac{n_i}{N})\, s_i^2} \quad .$$

The statistic $\underline{F}^*$ is approximately distributed as $\underline{F}$ with G-1 and

$$f = \frac{\left[\sum_{i=1}^{G} \left(1 - \frac{n_i}{N}\right) s_i^2 \right]^2}{\sum_{i=1}^{G} \frac{\left[\left(1 - \frac{n_i}{N}\right) s_i^2 \right]^2}{n_i - 1}}$$

degrees of freedom. In the case of two groups, both the Brown-Forsythe test and Welch (1951) APDF test are equivalent to the Welch (1938) APDF test.

Wilcox (1989) proposed the statistic

$$H_m = \sum_{i=1}^{G} w_i (\tilde{X}_i - \tilde{X})^2$$

where

$$w_i = \left[\frac{s_i^2}{n_i} \right]^{-1} \quad i=1,\ldots,G$$

$$w = \sum_{i=1}^{G} w_i$$

$$\tilde{X}_i = \frac{2X_{in_i}}{n_i(n_i+1)} + \frac{n_i-1}{n_i(n_i+1)} \bar{X}_i \quad i=1,\ldots,G$$

and

$$\tilde{X} = \sum_{i=1}^{G} \frac{w_i \tilde{X}_i}{w} \quad .$$

The statistic $\underline{H}_m$ is approximately distributed as chi-square with G-1 degrees of freedom.

The literature suggests the following conclusions about control of Type I error rates under heteroscedastic conditions by the ANOVA $\underline{F}$, Welch APDF, James first- and second-order,

Brown-Forsythe, Wilcox $\underline{H}$, and Wilcox $\underline{H}_m$ tests: (a) the performance of each of these alternatives to ANOVA $\underline{F}$ is superior to $\underline{F}$; (b) the Welch test outperforms the James first-order test; (c) the Welch and Brown-Forsythe tests are generally competitive with one another, however, the Welch test is preferred with data sampled from normal distributions while the Brown-Forsythe test is preferred with data sampled from skewed distributions; and (d) the Wilcox $\underline{H}_m$ and James second-order test outperform all of these other alternatives to ANOVA $\underline{F}$ under the greatest variety of conditions. Brown and Forsythe (1974) used Monte Carlo techniques to examine the ANOVA $\underline{F}$, Brown-Forsythe $\underline{F}^*$, Welch APDF, and James zero-order procedures when (a) equal- and unequal-sized samples were selected from normal populations; (b) G was 4, 6, or 10; (c) the ratio of the largest to the smallest sample size was 1, 1.9, or 3; (d) the ratio of the largest to the smallest standard deviation was 1 or 3; and (e) total sample size ranged between 16 and 200. For small sample sizes the critical value proposed by Welch is a better approximation to the true critical value than is that proposed by James. Both the Welch APDF test and Brown-Forsythe $\underline{F}^*$ test have τ near α under the inequality of variances.

Kohr and Games (1974) examined the ANOVA $\underline{F}$ test, Box test, and Welch APDF test when (a) equal- and unequal-sized samples were selected from normal populations; (b) G was 4; (c) the ratio of the largest to the smallest sample size was

1, 1.5, or 2.8; (d) the ratio of the largest to the smallest standard deviation was 1, 2.0, $\sqrt{7}$, $\sqrt{10}$, or $\sqrt{13}$; and (e) total sample size ranged between 32 and 34. The best control of Type I error rates was demonstrated by the Welch APDF test. Kohr and Games concluded the Welch test may be used with confidence with the unequal-sized samples and the heteroscedastic conditions examined in their study. Kohr and Games concluded the Welch test was slightly liberal under heteroscedastic conditions; however, this bias was trivial compared to the inflated error rates for the $\underline{F}$ test and Box test under comparable conditions. Levy (1978) examined the Welch test when data were sampled from either the uniform, chi-square, or exponential distributions and also found that, under heteroscedasticity, the Welch test can be liberal.

Dijkstra and Werter (1981) compared the James second-order, Welch APDF, and Brown-Forsythe tests when (a) equal- and unequal-sized samples were selected from normal populations; (b) G was 3, 4, or 6; (c) the ratio of the largest to the smallest sample size was 1, 2, or 2.5; (d) total sample size ranged between 12 and 90; and (e) the ratio of the largest to the smallest standard deviation was 1 or 3. Dijkstra and Werter concluded the James second-order test gave better control of Type I error rates than either the Brown-Forsythe $\underline{F}^*$ or Welch APDF test.

Clinch and Keselman (1982) studied the ANOVA $\underline{F}$, Welch APDF, and Brown-Forsythe $\underline{F}^*$ tests using Monte Carlo methods

when (a) equal- and unequal-sized samples were selected from normal distributions, chi-square distributions with two degrees of freedom, or t distributions with five degrees of freedom; (b) G was 4; (c) the ratio of the largest to the smallest sample size was 1 or 3; (d) total sample size was 48 or 144; and (e) variances were either homoscedastic or heteroscedastic. The ANOVA F test was most affected by assumption violations. Type I error rates of the Welch test were above α, especially in the negative case. The F^* test provided the best Type I error control in that it generally only became nonrobust with extreme heteroscedasticity. Although both the Brown-Forsythe test and Welch test were liberal with skewed distributions, the tendency was stronger for the Welch test.

Tomarken and Serlin (1986) examined six tests including the ANOVA F test, Brown-Forsythe test, and Welch APDF test when (a) equal- and unequal-sized samples were selected from normal populations; (b) G was 3 or 4; (c) the ratio of the largest to the smallest sample size was 1 or 3; (d) total sample size ranged between 36 and 80; and (e) the ratio of the largest to smallest standard deviation was 1, 6, or 12. Tomarken and Serlin found that the Brown-Forsythe F^* test, though generally acceptable, was at least slightly liberal whether sample sizes were equal or directly or inversely paired with variances.

Wilcox, Charlin, and Thompson (1986) examined Monte Carlo results on the robustness of the ANOVA $\underline{F}$, Brown-Forsythe $\underline{F}^*$, and the Welch APDF test when (a) equal- and unequal-sized samples were selected from normal populations; (b) G was 2, 4, or 6; (c) the ratio of the largest to the smallest sample size was 1, 1.9, 3, 3.3, or 4.2; (d) total sample size ranged between 22 and 95; and (e) the ratio of the largest to smallest standard deviation was 1 or 4. Wilcox, Charlin, and Thompson gave practical situations where both the Welch and $\underline{F}^*$ tests may not provide adequate control over Type I error rates. For equal variances but unequal-sized samples, the Welch test should be avoided in favor of the $\underline{F}^*$ test but for unequal-sized samples and possibly unequal variances, the Welch test was preferred to the $\underline{F}^*$ test.

Wilcox (1988) proposed $\underline{H}$, a competitor to the Brown-Forsythe $\underline{F}^*$, Welch APDF, and James second-order test. Simulated equal- and unequal-sized samples were selected where (a) distributions were either normal, light-tailed symmetric, heavy-tailed symmetric, medium-tailed asymmetric, or exponential-like; (b) G was 4, 6, or 10; (c) the ratio of the largest to the smallest sample size was 1, 1.8, 2.5, 3.7, or 5; (d) total sample size ranged between 44 and 100; and (e) the ratio of the largest to the smallest standard deviation was 1, 4, 5, 6, or 9. These simulations indicated that under normality the new procedure always gave the experimenter as good or better control over the probability of a Type I error

than did the $\underline{F}^*$ test or Welch APDF test. Wilcox showed that, under normality, James's second-order test and Wilcox's test have τ much closer to α than the Welch or Brown-Forsythe tests. The Wilcox test gave conservative results provided $n_i \geq 10$ $(i=1,\ldots,G)$. Wilcox's results indicate the $\underline{H}$ procedure has a Type I error rate that is similar to James's second-order method, regardless of the degree of heteroscedasticity. Although computationally more tedious, Wilcox recommended James's second-order procedure for general use.

Wilcox (1989) proposed $\underline{H}_m$, an improvement to Wilcox's (1988) $\underline{H}$ method, designed to be more comparable in power to James's second-order test. Wilcox compared James's second-order test with $\underline{H}_m$ when (a) data were sampled from normal populations; (b) G was 4 or 6; (c) the ratio of the largest to the smallest sample size was 1, 2.5, 2.7, or 5; (d) total sample size ranged between 44 and 121; and (e) the ratio of the largest to the smallest standard deviation was 1, 4, or 6. Wilcox's results indicate that when applied to normal heteroscedastic data, $\underline{H}_m$ has τ near α and slightly less power than James's second-order test. The main advantage of the improved Wilcox procedure is that it is much easier to use than James's second order test, and it is easily extended to higher way designs.

Oshima and Algina (in press) studied Type I error rates for the Brown-Forsythe test, James's second-order test, Welch's APDF test, and Wilcox's $\underline{H}_m$ test for 155 conditions.

These conditions were obtained by crossing the 31 conditions defined by sample sizes and standard deviations in the Wilcox (1988) study with five distributions--normal, uniform, $\underline{t}(5)$, beta(1.5,8.5), and exponential. The James second-order test and Wilcox $\underline{H}_m$ test were both affected by non-normality. When samples were selected from symmetric non-normal distributions both James's second-order test and Wilcox's $\underline{H}_m$ test maintained τ near α. When the tests were applied to data sampled from asymmetric distributions, $|\tau-\alpha|$ increased. Further, as the degree of asymmetry increased, $|\tau-\alpha|$ tended to increase. The Brown-Forsythe test outperformed the Wilcox $\underline{H}_m$ test and James's second-order test under some conditions, however, the reverse held under other conditions. Oshima and Algina concluded (a) the Wilcox $\underline{H}_m$ test and James's second-order test were preferable to the Brown-Forsythe test, (b) James's second-order test was recommended for data sampled from a symmetric distribution, and (c) Wilcox's $\underline{H}_m$ test was recommended for data sampled from a moderately skewed distribution.

In summary, when data are sampled from a normal distribution, the Wilcox $\underline{H}_m$ test and James second-order test have better control of Type I error rates, particularly as the degree of heteroscedasticity gets large. All of these alternatives to the ANOVA $\underline{F}$ are affected by skewed data but there is some evidence the Brown-Forsythe $\underline{F}^*$ test and Wilcox $\underline{H}_m$ test are less affected.

Hotelling's T^2 Test

Hotelling's (1931) T^2 is a test of the equality of two population mean vectors when independent random samples are selected from two populations which are distributed multivariate normal and have equal dispersion matrices. The test statistic is given by

$$T^2 = \frac{n_1 n_2}{n_1+n_2}(\bar{x}_1 - \bar{x}_2)' \boldsymbol{S}^{-1} (\bar{x}_1 - \bar{x}_2)$$

where

$$\boldsymbol{S} = \frac{(n_1-1)\boldsymbol{S}_1 + (n_2-1)\boldsymbol{S}_2}{n_1+n_2-2} \quad .$$

Hotelling demonstrated the transformation

$$\frac{n_1+n_2-p-1}{(n_1+n_2-2)p} T^2$$

has an F distribution with p and n_1+n_2-p-1 degrees of freedom.

The sensitivity of Hotelling's T^2 to violations of the assumption of homoscedasticity is well documented. This has been investigated both analytically (Ito & Schull, 1964) and empirically (Algina & Oshima, 1990; Hakstian, Roed, & Lind, 1979; Holloway & Dunn, 1967; Hopkins & Clay, 1963). Ito and Schull (1964) investigated the large sample properties of T^2 in the presence of unequal dispersion matrices Σ_1 and Σ_2. Ito and Schull showed that in the case of two very large equal-sized samples, T^2 is well behaved even when the dispersion matrices are not equal and that in the case of two samples of nearly equal size, the test is not affected by moderate

inequality of dispersion matrices provided the samples are very large. However, if the two samples are of unequal size, quite a large effect occurs on the level of significance from even moderate variations. Ito and Schull indicated that, asymptotically, with fixed $n_1/(n_1+n_2) > 0.5$ and for equal eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$, $\tau < \alpha$ when the eigenvalues are greater than one and $\tau > \alpha$ when the eigenvalues are less than one.

Hopkins and Clay (1963) examined distributions of Hotelling's $\underline{T}^2$ with sample sizes of 5, 10, and 20 selected from either (a) bivariate normal populations with zero means, dispersion matrices of the form $\sigma_i^2\underline{I}$ (i=1,2), where σ_2/σ_1 was 1, 1.6, or 3.2; or (b) circular bivariate symmetrical leptokurtic populations with zero means, equal variances, and β_2-3 was 3.2 or 6. Hopkins and Clay reported $\underline{T}^2$ is robust to violations of homoscedasticity when $n_1=n_2 \geq 10$ but that this robustness does not extend to disparate sample sizes. Hopkins and Clay reported that upper tail frequencies of the distribution of Hotelling's $\underline{T}^2$ for $n_i \geq 10$ (i=1,2) are not substantially affected by moderate degrees of symmetrical leptokurtosis.

Holloway and Dunn (1967) examined the robustness of Hotelling's $\underline{T}^2$ to violations of the homoscedasticity assumption when (a) equal- and unequal-sized samples were selected from multivariate normal distributions; (b) p was 1, 2, 3, 5, 7, or 10; (c) total sample size ranged between 10 and 200; (d) $n_1/(n_1+n_2)$ was .3, .4, .5, .6, or .7; and (e) the

eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$ were 3 or 10. Holloway and Dunn found equal-sized samples help in keeping τ close to α. Further, Holloway and Dunn found that for large equal-sized samples, control of Type I error rates depends on the number of dependent variables (p). For example, when $n_i = 50$ ($i=1,2$) and all the eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1} = 10$, τ is near α for $p = 2$ and $p = 3$, but τ markedly departs from α when $p = 7$ or $p = 10$. Holloway and Dunn found that generally as the number of dependent variables increases, or as the sample size decreases, τ increases.

Hakstian, Roed, and Lind (1979) obtained empirical sampling distributions of Hotelling's $\underline{T}^2$ when (a) equal- and unequal-sized samples were selected from multivariate normal populations; (b) p was 2, 6, or 10; (c) $(n_1+n_2)/2$ was 3 or 10; (d) n_1/n_2 was 1, 2, or 5; and (e) dispersion matrices were of the form $\mathbf{I}$ and $\mathbf{D}$, where $\mathbf{D}$ was $\mathbf{I}$, $d^2\mathbf{I}$, or $\text{diag}\{1,1,\ldots,1,d^2,d^2,\ldots,d^2\}$ ($d = 1$, 1.2, or 1.5). Hakstian, Roed, and Lind found that for equal-sized samples, the $\underline{T}^2$ procedure is generally robust. With unequal-sized samples, $\underline{T}^2$ was shown to become increasingly less robust as dispersion heteroscedasticity and the number of independent variables increase. Consequentially, Hakstian, Roed, and Lind argued against the use of $\underline{T}^2$ in the negative condition and for cautious use in the positive condition.

Algina and Oshima (1990) studied Hotelling's $\underline{T}^2$ where (a) p was 2, 6, or 10; (b) the ratio of total sample size to

number of dependent variables was 6, 10, or 20; and (c) for the majority of conditions $\mathbf{\Sigma}_2 = d^2\mathbf{\Sigma}_1$ (d = 1.5, 2.0, 2.5, or 3.0). Algina and Oshima found that even with a small sample size ratio, the $\underline{T}^2$ procedure can be seriously nonrobust. For example, with p = 2 and $\mathbf{\Sigma}_2 = 2.25\mathbf{\Sigma}_1$, a sample size ratio as small as 1.1:1 can produce unacceptable Type I error rates. Algina and Oshima also confirmed earlier findings that Hotelling's $\underline{T}^2$ test became less robust as the number of dependent variables and degree of heteroscedasticity increased.

In summary, Hotelling's $\underline{T}^2$ test is not robust to violations of the assumption of homoscedasticity even when there are equal-sized samples, especially if the ratio of total sample size to number of dependent variables is small. When the larger sample is selected from the population with the larger dispersion matrix, $\tau < \alpha$. When the larger sample is selected from the population with the smaller dispersion matrix, $\tau > \alpha$. These tendencies increase with the inequality of the size of the two samples, the degree of heteroscedasticity, and the number of dependent variables.

Therefore, the behavior of Hotelling's $\underline{T}^2$ test is similar to the independent samples $\underline{t}$ test under violations of the assumption of homoscedasticity. Hence, it is desirable to examine robust alternatives that do not require this basic assumption of the Hotelling's $\underline{T}^2$ procedure.

Alternatives to the Hotelling's T^2 Test

A number of tests have been developed to test the hypothesis $H_0\colon \mu_1 = \mu_2$ in a situation in which $\Sigma_1 \neq \Sigma_2$. Alternatives to the Hotelling T^2 procedure that do not assume equality of the two population dispersion matrices include James's (1954) first- and second-order tests, Yao's (1965) test, and Johansen's (1980) test. Differing only in their critical values, all four tests use the test statistic

$$T_v^2 = (\bar{x}_1 - \bar{x}_2)' \left[\frac{S_1}{n_1} + \frac{S_2}{n_2} \right]^{-1} (\bar{x}_1 - \bar{x}_2)$$

where $\bar{\mathbf{x}}_1$ and $\mathbf{S}_i$ are respectively the sample mean vector and sample dispersion matrix for the ith sample (i=1,2).

The literature suggests the following conclusions about control of Type I error rates under heteroscedastic conditions by Hotelling's T^2 test, James's first- and second-order tests, Yao's test, and Johansen's test: (a) Yao's test, James's second-order test, and Johansen's test are superior to James's first-order test; and (b) all of these alternatives to Hotelling's T^2 are sensitive to data sampled from skewed populations.

Yao (1965) conducted a Monte Carlo study to compare Type I error rates between the James first-order test and the Yao test when (a) equal- and unequal-sized samples were selected, (b) p was 2, (c) the ratio of total sample size to number of dependent variables was 10 or 13, and (d) dispersion matrices

were unequal. Although both procedures have τ near α under heteroscedasticity, Yao's test was superior to James's test.

Algina and Tang (1988) examined the performance of Hotelling's $\underline{T}^2$, James's first-order test, and Yao's test when (a) p was 2, 6, or 10; (b) N:p was 6, 10, or 20; (c) the ratio of the largest to the smallest sample size was 1, 1.25, 1.5, 2, 3, 4, or 5; and (d) the dispersion matrices were of the form **I** and **D**, where **D** was $d^2\mathbf{I}$ (d = 1.5, 2.0, 2.5, or 3.0), diag{3,1,1,...,1}, diag{3,3,...,3,1,1,...,1}, diag{1/3,3,3,...,3}, or diag{1/3,1/3,...,1/3,3,3,...,3}. Algina and Tang confirmed the superiority of Yao's test. For $10 \le N{:}p \le 20$, Yao's test produced appropriate Type I error rates when $p \le 10$, $n_1{:}n_2 \le 2{:}1$, and $d \le 3$. For N:p = 20, appropriate error rates occurred when $n_1{:}n_2 \le 5{:}1$ and $d \le 3$. This applied for both the specific cases where one dispersion matrix was a multiple of the second ($\mathbf{\Sigma}_2 = d^2\mathbf{\Sigma}_1$) and in more complex cases of heteroscedasticity. When N:p = 6 and $\mathbf{\Sigma}_2 = d^2\mathbf{\Sigma}_1$, Algina and Tang found Yao's test to be liberal.

Algina, Oshima, and Tang (1991) studied Type I error rates for James's first- and second-order, Yao's, and Johansen's tests for various conditions defined by the degree of heteroscedasticity and non-normality (uniform, Laplace, $\underline{t}(5)$, beta(5,1.5), exponential, and lognormal distributions). The study indicated these four alternatives to Hotelling's $\underline{T}^2$ may not be robust, when the sampled distributions have heteroscedastic dispersion matrices, are skewed, and have

positive kurtosis. Although the four procedures were seriously nonrobust with exponential and lognormal distributions, they were fairly robust with the remaining distributions. The performance of Yao's test, James's second-order test, and Johansen's test was slightly superior to the performance of James's first-order test. Algina, Oshima, and Tang indicate that Yao's test is also sensitive to skewness.

In summary Yao's test, James's second-order test, and Johansen's test work reasonably well under normality. Although all of these alternatives to Hotelling's $\underline{T}^2$ test have elevated Type I error rates with skewed data, Johansen's test has the practical advantages of (a) generalizing to G > 2, and (b) being relatively easy to compute.

MANOVA Criteria

The four basic multivariate analysis of variance (MANOVA) criteria are used to test the equality of G population mean vectors when independent random samples are selected from populations which are distributed multivariate normal and have equal dispersion matrices. Define

$$\boldsymbol{H} = \sum_{i=1}^{G} n_i (\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})'$$

and

$$\boldsymbol{E} = \sum_{i=1}^{G} (n_i - 1)\boldsymbol{S}_i \quad .$$

The basic MANOVA criteria are all functions of the eigenvalues of $\mathbf{HE}^{-1}$. Define τ_i to be the ith eigenvalue of $\mathbf{HE}^{-1}$ ($i=1,\ldots,s$), where $s = \min(p,G-1)$. Those criteria are:

1. Roy's (1945) largest root criterion

$$R = \frac{\tau_1}{1 + \tau_1} \quad ;$$

2. Hotelling-Lawley trace criterion (Hotelling, 1951; Lawley, 1938)

$$U = trace(\boldsymbol{HE}^{-1}) = \sum_{i=1}^{s} \tau_i \quad ;$$

3. Pillai-Bartlett trace criterion (Pillai, 1955; Bartlett, 1939)

$$V = trace[\boldsymbol{H}(\boldsymbol{H}+\boldsymbol{E})^{-1}] = \sum_{i=1}^{s} \frac{\tau_i}{1+\tau_i} \quad ;$$

and

4. Wilks's (1932) likelihood ratio criterion

$$L = \frac{|\boldsymbol{E}|}{|\boldsymbol{H}+\boldsymbol{E}|} = \prod_{i=1}^{s} \frac{\tau_i}{1+\tau_i} \quad .$$

Both analytic (Pillai & Sudjana, 1975) and empirical (Korin, 1972; Olson, 1974) investigations have been conducted on the robustness of MANOVA criteria with respect to violations of homoscedasticity. Pillai and Sudjana (1975) examined violations of homoscedasticity on the four basic MANOVA criteria. Although the generalizability of the study was limited by only examining equal-sized samples selected from two populations with unclear degrees of

heteroscedasticity, the results were consistent--modest departures from α for minor degrees of heteroscedasticity and more pronounced departures with greater heteroscedasticity.

Korin (1972) studied Roy's largest root criterion ($\underline{R}$), the Hotelling-Lawley trace criterion ($\underline{U}$), and Wilks's likelihood ratio criterion ($\underline{L}$) when (a) equal- and unequal-sized samples were selected from normal populations; (b) p was 2 or 4; (c) G was 3 or 6; (d) the ratio of total sample size to number of dependent variables was 8.25, 9, 12, 15.5, 18 or 33; and (e) dispersion matrices were of the form **I** or **D**, where **D** was $d^2\mathbf{I}$ or $2d^2\mathbf{I}$ (d = 1.5 or 10). For small samples, even when the sample sizes were all equal, dispersion heteroscedasticity produced Type I error rates greater than α. Korin reported the error rates for $\underline{R}$ were greater than those for $\underline{U}$ and $\underline{L}$.

Olson (1974) conducted a Monte Carlo study on the comparative robustness of six multivariate tests including the four basic MANOVA criteria ($\underline{R}$, $\underline{U}$, $\underline{L}$, $\underline{V}$) when (a) equal-sized samples were selected; (b) p was 2, 3, 6, or 10; (c) G was 2, 3, 6, or 10; (d) n_i was 5, 10, or 50 (i=1,...,G); and (e) dispersion matrices were of the form **I** or **D**, where **D** represented either a low or high degree of contamination. For the low degree of contamination, $\mathbf{D} = d^2\mathbf{I}$, whereas for the high degree of contamination, $\mathbf{D} = \mathrm{diag}\{pd^2-p+1,1,1,\ldots,1\}$ (d = 2, 3, or 6). Results indicated that for protection against non-normality and heteroscedasticity of dispersion matrices, $\underline{R}$

should be avoided, while V may be recommended as the most robust of the MANOVA tests. In terms of the magnitude of the departure of τ from α, the order was typically R > U > L > V. This tendency increased as the degree of heteroscedasticity increased. The departure of τ from α for R, U, and L increased with an increase in the number of dependent variables, however, the impact of p on V was not as well defined. Additionally, for R, U, and L, τ decreased as sample size increased except when G > 6. When G > 6, τ increased for all four basic MANOVA procedures, although the increase was least for V.

Stevens (1979) contested Olson's (1976) claim that V is superior to L and U for general use in multivariate analysis of variance because of greater robustness against unequal dispersion matrices. Stevens believed Olson's conclusions were tainted by using an example which had extreme subgroup variance differences, which occur very infrequently in practice. Stevens conceded V was the clear choice for diffuse structures, however, for concentrated noncentrality structures with dispersion heteroscedasticity, the actual Type I error rates for V, U, and L are very similar. Olson (1979) refuted Stevens's (1979) objections on practical grounds. The experimenter, faced with real data of unknown noncentrality and trying to follow Stevens's recommendation to use V for diffuse noncentrality and any of the V, U, or L statistics for concentrated noncentrality, must always choose V.

Alternatives to the MANOVA Criteria

A number of tests have been developed to test the hypothesis $H_o: \mu_1 = \mu_2 = \ldots = \mu_G$ in a situation in which $\Sigma_i \neq \Sigma_j$ (for at least one pair of i and j). James (1954) generalized James's (1951) series solutions and proposed the statistic

$$J = \sum_{i=1}^{G} (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})' \boldsymbol{W}_i (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})$$

where

$$\boldsymbol{W}_i = \left[\frac{\boldsymbol{S}_i}{n_i}\right]^{-1} \qquad i=1,\ldots,G$$

$$\boldsymbol{W} = \sum_{i=1}^{G} \boldsymbol{W}_i$$

$$\bar{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij} \qquad i=1,\ldots,G$$

and

$$\bar{\boldsymbol{x}} = \boldsymbol{W}^{-1} \sum_{i=1}^{G} \boldsymbol{W}_i \bar{\boldsymbol{x}}_i \quad .$$

The James (1954) zero-, first-, and second-order critical values parallel those developed by James (1951).

Johansen (1980) generalized the Welch (1951) test and proposed using the James (1954) test statistic $\underline{J}$ divided by

$$c = p(G-1) + 2A - \frac{6A}{p(G-1)+2}$$

where

$$A = \sum_{i=1}^{G} \frac{trace(\boldsymbol{I}-\boldsymbol{W}^{-1}\boldsymbol{W}_i)^2 + trace^2(\boldsymbol{I}-\boldsymbol{W}^{-1}\boldsymbol{W}_i)}{2(n_i-1)} .$$

The critical value for the Johansen test is a fractile of an F distribution with p(G-1) and p(G-1)[p(G-1)+2]/(3A) degrees of freedom.

The literature suggests the following conclusions about control of Type I error rates when sampling from multivariate normal populations under heteroscedastic conditions by the four basic MANOVA criteria, James's first- and second-order tests, and Johansen's test: (a) the Pillai-Bartlett trace criterion is the most robust of the four basic MANOVA criteria; and (b) with unequal-sized samples, Johansen's test and James's second-order test outperform the Pillai-Bartlett trace criterion and James's first-order test.

Ito (1969) analytically examined Type I error rates for James's zero-order test and showed $\tau > \alpha$. Ito showed $|\tau-\alpha|$ increased as the variation in the sample sizes, degree of heteroscedasticity, and number of dependent variables increased, whereas $|\tau-\alpha|$ decreased as the total sample size increased.

Tang (1989) studied the Pillai-Bartlett trace criterion, James's first- and second-order tests, and Johansen's test when (a) equal- and unequal-sized samples were selected from multivariate normal populations; (b) p was 3 or 6; (c) G was 3; (d) the ratio of the largest to the smallest sample size ratio was 1, 1.3, or 2; (e) the ratio of total sample size to

number of dependent variables was 10, 15, or 20; and (f) dispersion matrices were either of the form **I** or **D**, where **D** was $d^2\mathbf{I}$, diag$\{1,d^2,d^2\}$, or diag$\{1/d^2,d^2,d^2\}$ for p=3 or **D** was $d^2\mathbf{I}$, diag$\{1,1,1,d^2,d^2,d^2\}$,or diag$\{1/d^2,1/d^2,1/d^2,d^2,d^2,d^2\}$ for p=6 (d = $\sqrt{1.5}$ or 3). Results of the study indicate when (a) sample sizes are unequal and (b) dispersion matrices are unequal, Johansen's test and James's second-order test perform better than the Pillai-Bartlett trace criterion and James first-order test. While both Johansen's test and James's second-order test tended to have Type I error rates reasonably near α, Johansen's test was slightly liberal whereas James's second-order test was slightly conservative. Additionally, the ratio of total sample size to number of dependent variables (N:p) has a strong impact on the performance of the tests. Generally, as N:p increases, the test becomes more robust.

In summary, the Pillai-Bartlett trace criterion appears to be the most robust of the four basic MANOVA criteria to violations of the assumption of dispersion homoscedasticity. In controlling type I error rates, the Johansen test and James second-order test are more effective than either the Pillai-Bartlett trace criterion or James first-order test. Finally, the Johansen test has the practical advantage of being less computationally intensive than the James second-order test.

CHAPTER 3
METHODOLOGY

In this chapter, the development of the test statistics, the design, and the simulation procedure are described. The test statistics extend the work of Brown and Forsythe (1974) and Wilcox (1989). The design is based upon a review of relevant literature and upon the consideration that the experimental conditions used in the simulation should be similar to those found in educational research.

Development of Test Statistics

Brown-Forsythe Generalizations

To test H_0: $\mu_1 = \mu_2 = \ldots = \mu_G$ Brown and Forsythe (1974) proposed the statistic

$$F^* = \frac{\sum_{i=1}^{G} n_i (\bar{x}_{i.} - \bar{x}_{..})^2}{\sum_{i=1}^{G} (1 - \frac{n_i}{N}) s_i^2} .$$

The statistic $\underline{F}^*$ is approximately distributed as $\underline{F}$ with G-1 and f degrees of freedom, where

$$f = \frac{[\sum_{i=1}^{G} (1 - \frac{n_i}{N}) s_i^2]^2}{\sum_{i=1}^{G} \frac{[(1 - \frac{n_i}{N}) s_i^2]^2}{n_i - 1}} .$$

Suppose $\bar{\mathbf{x}}_1,\ldots,\bar{\mathbf{x}}_G$ are p-dimensional sample mean vectors and $\mathbf{S}_1,\ldots,\mathbf{S}_G$ are p-dimensional dispersion matrices of independent random samples of sizes $n_1,\ldots,n_G$, respectively, from G multivariate normal distributions $N_p(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1),\ldots,N_p(\boldsymbol{\mu}_G,\boldsymbol{\Sigma}_G)$. To extend the Brown-Forsythe statistic to the multivariate setting, replace means by corresponding mean vectors and replace variances by their corresponding dispersion matrices. Define

$$\boldsymbol{H} = \sum_{i=1}^{G} n_i (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})'$$

and

$$\boldsymbol{M} = \sum_{i=1}^{G} (1 - \frac{n_i}{N}) \boldsymbol{S}_i \ .$$

The $\mathbf{S}_i$ $(i=1,\ldots,G)$ are distributed independently as Wishart $W_p(n_i,\boldsymbol{\Sigma}_i)$ and $\mathbf{M}$ is said to have a sum of Wisharts distribution, denoted as $\mathbf{M} \sim SW(n_1,\ldots,n_G;(1 - n_1/N)\boldsymbol{\Sigma}_1,\ldots,(1 - n_G/N)\boldsymbol{\Sigma}_G)$. Nel and van der Merwe (1986) have generalized Satterthwaite's (1946) results and approximated the sum of Wisharts distribution by $\mathbf{Z} \sim W_p(f,\boldsymbol{\Sigma})$. Applying the Nel and van der Merwe results to $\mathbf{M}$, the quantity f is the approximate degrees of freedom of $\mathbf{M}$ and is given by

$$f = \frac{trace^2[\sum_{i=1}^{G} c_i\boldsymbol{\Sigma}_i] + trace[\sum_{i=1}^{G} c_i\boldsymbol{\Sigma}_i]^2}{\sum_{i=1}^{G} \frac{1}{n_i - 1}\{trace^2[c_i\boldsymbol{\Sigma}_i] + trace[c_i\boldsymbol{\Sigma}_i]^2\}}$$

where

$$c_i = 1 - \frac{n_i}{N} \quad .$$

The problem is to construct test statistics and determine critical values. The approach used in this study is to construct test statistics analogous to those developed by Lawley-Hotelling (U), Pillai-Bartlett (V), and Wilks (L). Define

$$\boldsymbol{H} = \sum_{i=1}^{G} n_i (\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}}) (\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})'$$

and

$$\boldsymbol{E} = \sum_{i=1}^{G} (n_i - 1) \boldsymbol{S}_i \quad .$$

Then the test statistics for the Hotelling-Lawley trace criterion, the Pillai-Bartlett trace criterion, and the Wilks likelihood ratio criterion are, respectively

$$U = trace[\, \boldsymbol{HE}^{-1} \,]$$

$$V = trace[\boldsymbol{H}(\boldsymbol{H} + \boldsymbol{E})^{-1}]$$

and

$$L = \frac{|\boldsymbol{E}|}{|\boldsymbol{H} + \boldsymbol{E}|} \quad .$$

Approximate F transformations can be used with each of these test statistics. Define the following variables:

p = number of independent variables

h = G - 1 (the degrees of freedom for the multivariate analog to sums of squares between groups)

s = min(p,h)

e = N - G (the degrees of freedom for the multivariate analog to sums of squares within groups)

m = .5(|p - h| - 1)

n = .5(e - p - 1)

For the Hotelling-Lawley criterion, the transformations developed by Hughes and Saw (1972) and McKeon (1974) respectively are given by

$$F_U^{(1)} = \frac{2(sn+1)}{s(2m+s+1)}\frac{U}{s} \sim F_{s(2m+s+1),2(sn+1)}$$

and

$$F_U^{(2)} = \frac{2n}{a-2}\frac{a}{ph} U \sim F_{ph,a}$$

where

$$a = 4 + \frac{ph + 2}{b - 1}$$

and

$$b = \frac{(2n + h)(2n + p)}{2(n - 1)(2n + 1)} .$$

For the Pillai-Bartlett criterion the SAS (1985, p.12) transformation is given by

$$F_V = \frac{2n+s+1}{2m+s+1}\frac{V}{s - V} \sim F_{s(2m+s+1),s(2n+s+1)} .$$

For the Wilks criterion, the Rao (1952, p.262) transformation is given by

$$F_L = \frac{1 - L^{\frac{1}{t}}}{L^{\frac{1}{t}}}\frac{rt - 2q}{ph} \sim F_{ph,rt - 2q}$$

where

$$q = \frac{ph - 2}{4}$$

$$t = \sqrt{\frac{p^2h^2 - 4}{p^2 + h^2 - 5}} \quad if\ p^2 + h^2 - 5 > 0$$

$$t = 1,\ otherwise$$

and

$$r = e - \frac{p + h + 1}{2} \quad .$$

<u>Scale of the measures of between and within group variability</u>. Consider for the univariate (p=1) case the denominator of the Brown-Forsythe statistic ($\underline{F}^*$)

$$M = \sum_{i=1}^{G} (1 - \frac{n_i}{N}) s_i^2$$

$$= \sum_{i=1}^{G} s_i^2 - \sum_{i=1}^{G} \frac{n_i s_i^2}{N}$$

$$= G\ [\ \frac{1}{G}\sum_{i=1}^{G} s_i^2\ ] - \sum_{i=1}^{G} \frac{n_i s_i^2}{N}$$

$$= G\bar{s}_.^2 - \tilde{s}_.^2$$

Here $\bar{s}_.^2$ is the arithmetic average of the G sample variances and $\tilde{s}_.^2$ is the average of the G sample variances weighted by their respective sample sizes. Because both are approaches to approximating the average dispersion, M roughly represents a mean square within group (MSWG) multiplied by the degrees of

freedom for the sum of squares between groups. Because the numerator of $\underline{F}^*$ is the between group sums of squares, the Brown-Forsythe statistic is in the metric of the ratio of two mean squares. Now the MANOVA criteria are in the metric of the ratio of two sum of squares. Consider the common MANOVA criteria in the univariate setting. For Hotelling-Lawley, Pillai-Bartlett, and Wilks respectively, $\underline{U}$ = SSBG/SSWG, $\underline{V}$ = SSBG/(SSBG+SSWG), and $\underline{L}$ = SSWG/(SSBG+SSWG). In each case the test statistics are functions of the sum of squares rather than mean squares. Hence, in order to use criteria analogous to $\underline{U}$, $\underline{V}$, and $\underline{L}$, $\mathbf{E}$ must be replaced by $(f/h)\mathbf{M}$.

Let τ_i^*, i=1,...,s be the ith eigenvalue of the characteristic equation $|\mathbf{H} - \tau_i^*(f/h)\mathbf{M}|=0$. One statistic to consider would be analogous to Roy's largest root criterion (1945) $\tau_1^*/(1 + \tau_1^*)$ where $\tau_1^* > \ldots > \tau_s^*$. Of the four basic MANOVA criteria, Roy's largest root criterion is the most affected by heteroscedasticity (Olson, 1974, 1976, 1979; Stevens, 1979). Consequentially, τ_1^* will be omitted. The Lawley-Hotelling trace (Hotelling, 1951; Lawley, 1938) is based upon the same characteristic equation as Roy's largest root criterion (1945). In this case, the analogous statistic $\underline{U}^*$ = $\mathrm{trace}\{\mathbf{H}[(f/h)\mathbf{M}]^{-1}\} = \Sigma\ \tau_i^*$ provides one of the test statistics of interest.

Let θ_i, i=1,...,s denote the ith eigenvalue of the characteristic equation $|\mathbf{H} - \theta_i[\mathbf{H}+(f/h)\mathbf{M}]^{-1}|=0$. Then the statistic analogous to the Pillai-Bartlett trace

(Bartlett,1939; Pillai, 1955) $\underline{V}^* = \text{trace}\{\mathbf{H}[\mathbf{H}+(f/h)\mathbf{M}]^{-1}\} = \Sigma\, \theta_i$ provides another test statistic of interest.

Similarly, if δ_i, i=1,...,s is the ith eigenvalue of the characteristic equation $|(f/h)\mathbf{M} - \delta_i(\mathbf{H}+(f/h)\mathbf{M})|=0$, then the analogous Wilks (1932) criterion is defined $\underline{L}^* = |(f/h)\mathbf{M}|/|\mathbf{H}+(f/h)\mathbf{M}| = \Pi\, \delta_i$.

To conduct hypothesis testing, approximate $\underline{F}$ transformations were used with each of these analogous test statistics, replacing N-G, the degrees of freedom for $\mathbf{S} = (N-G)^{-1}\mathbf{E}$, by f, the approximate degrees of freedom for $\mathbf{M}$. Thus, the variables are defined as follows:

p = number of independent variables

h = G - 1 (the degrees of freedom for the multivariate analog to sums of squares between groups)

s = min(p,h)

$$f = \frac{trace^2[\sum_{i=1}^{G} c_i \boldsymbol{S_i}] + trace[\sum_{i=1}^{G} c_i \boldsymbol{S_i}]^2}{\sum_{i=1}^{G} \frac{1}{n_i - 1}\{trace^2[c_i\boldsymbol{S_i}] + trace[c_i\boldsymbol{S_i}]^2\}}$$

where

$$c_i = 1 - \frac{n_i}{N} \qquad i=1,...=G$$

$$N = \sum_{i=1}^{G} n_i$$

m = .5(|p - h| - 1)

n^*= .5(f - p - 1) .

For the modified Hotelling-Lawley criterion, the Hughes and Saw (1972) and McKeon (1974) transformations respectively are now given by

$$F_{U^*}^{(1)} = \frac{2(sn^*+1)}{s(2m+s+1)} \frac{U^*}{s} \sim F_{s(2m+s+1),2(sn^*+1)}$$

and

$$F_{U^*}^{(2)} = \frac{2n^*}{a^*-2} \frac{a^*}{ph} U^* \sim F_{ph,a^*}$$

where

$$a^* = 4 + \frac{ph+2}{b^*-1}$$

and

$$b^* = \frac{(2n^*+h)(2n^*+p)}{2(n^*-1)(2n^*+1)} \quad .$$

For the modified Pillai-Bartlett criterion the SAS (1985, p.12) transformation is now given by

$$F_{V^*} = \frac{2n^*+s+1}{2m+s+1} \frac{V^*}{s-V^*} \sim F_{s(2m+s+1),s(2n^*+s+1)} \quad .$$

For the modified Wilks criterion, the Rao (1952, p.262) transformation is now given by

$$F_{L^*} = \frac{1-L^{*\frac{1}{t}}}{L^{*\frac{1}{t}}} \frac{r^*t-2q}{ph} \sim F_{ph,r^*t-2q}$$

where

$$q = \frac{ph-2}{4}$$

$$t = \sqrt{\frac{p^2h^2-4}{p^2+h^2-5}} \quad if \ p^2+h^2-5 > 0$$

otherwise

$$t = 1,$$

and

$$r^{*} = f - \frac{p + h + 1}{2} .$$

<u>Equality of expectation of the measures of between and within group dispersion</u>. The Brown-Forsythe statistic was constructed so that, under the null hypothesis, the expectations of the numerator and denominator are equal. To show the proposed multivariate generalization of the Brown-Forsythe statistic possesses the analogous property (that is, E(**H**)=E(**M**), assuming H_0: $\mu_1 = \ldots = \mu_G$ is true) the following results are useful:

1. $E(\overline{\boldsymbol{x}}_i) = \boldsymbol{\mu}_i = \boldsymbol{\mu}$.

2. $E(\overline{\boldsymbol{x}}) = \boldsymbol{\mu}$.

3. $E(\overline{\boldsymbol{x}}\,\overline{\boldsymbol{x}}') = var(\overline{\boldsymbol{x}}) + \boldsymbol{\mu}\,\boldsymbol{\mu}'$.

4. $E[(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu})(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu})'] = var(\overline{\boldsymbol{x}}_i) = \frac{1}{n_i}\boldsymbol{\Sigma}_i$.

5. $E[(\overline{\boldsymbol{x}} - \boldsymbol{\mu})(\overline{\boldsymbol{x}} - \boldsymbol{\mu})'] = var(\overline{\boldsymbol{x}}) = \frac{1}{n_.}\sum_{i=1}^{G}\frac{n_i\boldsymbol{\Sigma}_i}{n_.}$.

Using results 1-5, E(**M**) is given by

$$E(\boldsymbol{M}) = E[\sum_{i=1}^{G}(1 - \frac{n_i}{n_.})\boldsymbol{S}_i] ,$$

$$= \sum_{i=1}^{G}(1 - \frac{n_i}{n_.})E[\boldsymbol{S}_i] ,$$

$$= \sum_{i=1}^{G} (1 - \frac{n_i}{n_.}) \boldsymbol{\Sigma}_i \quad .$$

Similarly, using results 1-5, E(**H**) is given by

$$E(\boldsymbol{H}) = E[\sum_{i=1}^{G} n_i (\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}}) (\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})'] \quad ,$$

$$= \sum_{i=1}^{G} n_i \; E(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu}) (\overline{\boldsymbol{x}}_i - \boldsymbol{\mu})' + \sum_{i=1}^{G} n_i \; E(\overline{\boldsymbol{x}} - \boldsymbol{\mu}) (\overline{\boldsymbol{x}} - \boldsymbol{\mu})' \quad ,$$

$$= \sum_{i=1}^{G} n_i \; var(\overline{\boldsymbol{x}}_i) + \sum_{i=1}^{G} n_i \; var(\overline{\boldsymbol{x}})$$

$$- \; E\{ \sum_{i=1}^{G} n_i [\; \overline{\boldsymbol{x}}_i \, \overline{\boldsymbol{x}}' - \overline{\boldsymbol{x}}_i \boldsymbol{\mu}' - \boldsymbol{\mu} \overline{\boldsymbol{x}}' + \boldsymbol{\mu} \; \boldsymbol{\mu}' \;] \}$$

$$- \; E\{ \sum_{i=1}^{G} n_i [\overline{\boldsymbol{x}} \, \overline{\boldsymbol{x}}_i' - \overline{\boldsymbol{x}} \, \boldsymbol{\mu}' - \boldsymbol{\mu} \; \overline{\boldsymbol{x}}_i' + \boldsymbol{\mu} \; \boldsymbol{\mu}'] \} \quad ,$$

$$= \sum_{i=1}^{G} n_i (\frac{1}{n_i} \boldsymbol{\Sigma}_i) + \sum_{i=1}^{G} n_i [\; \frac{1}{n_.^2} \sum_{i=1}^{G} n_i \boldsymbol{\Sigma}_i \;]$$

$$-n_. \; E[\; \overline{\boldsymbol{x}} \, \overline{\boldsymbol{x}}' \;] + n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' + n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' - n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}'$$

$$-n_. \; E[\; \overline{\boldsymbol{x}} \, \overline{\boldsymbol{x}}' \;] + n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' + n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' - n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' \quad ,$$

$$= \sum_{i=1}^{G} \boldsymbol{\Sigma}_i + \frac{1}{n_.} \sum_{i=1}^{G} n_i \boldsymbol{\Sigma}_i - 2 n_. [\; var(\overline{\boldsymbol{x}}) + \boldsymbol{\mu} \; \boldsymbol{\mu}' \;] + 2 \; n_. \; \boldsymbol{\mu} \; \boldsymbol{\mu}' \quad ,$$

$$= \sum_{i=1}^{G} \boldsymbol{\Sigma}_i + \sum_{i=1}^{G} \frac{n_i \boldsymbol{\Sigma}_i}{n_.}$$

$$- 2n_. [\frac{1}{n_.} \sum_{i=1}^{G} \frac{n_i \boldsymbol{\Sigma}_i}{n_.} + \boldsymbol{\mu}\boldsymbol{\mu}'] + 2n_. \boldsymbol{\mu}\boldsymbol{\mu}' ,$$

$$= \sum_{i=1}^{G} \boldsymbol{\Sigma}_i + \sum_{i=1}^{G} \frac{n_i \boldsymbol{\Sigma}_i}{n_.} - 2 \sum_{i=1}^{G} \frac{n_i \boldsymbol{\Sigma}_i}{n_.} - 2n_. \boldsymbol{\mu}\boldsymbol{\mu}' + 2n_. \boldsymbol{\mu}\boldsymbol{\mu}' ,$$

$$= \sum_{i=1}^{G} \boldsymbol{\Sigma}_i - \sum_{i=1}^{G} \frac{n_i \boldsymbol{\Sigma}_i}{n_.} ,$$

$$= \sum_{i=1}^{G} (1 - \frac{n_i}{n_.}) \boldsymbol{\Sigma}_i .$$

Hence, $E(\mathbf{H}) = E(\mathbf{M})$.

Thus the modified Brown-Forsythe generalizations parallel the basic MANOVA criteria in terms of the measure of between group dispersion, the measure of within group dispersion, the metric of between and within group dispersion, and the equality of the expectation of the measures of between and within group dispersion.

Wilcox Generalization

To test H_0: $\mu_1 = \mu_2 = \ldots = \mu_G$ Wilcox (1989) proposed using the test statistic

$$H_m = \sum_{i=1}^{G} w_i (\tilde{X}_i - \tilde{X})^2 ,$$

where $\underline{H}_m$ is approximately distributed chi-square with G-1 degrees of freedom. To extend this to the multivariate setting, replace w_i with $\mathbf{W}_i$, $\tilde{x}_i$ with $\tilde{\mathbf{x}}_i$, $\tilde{x}$ with $\tilde{\mathbf{x}}$, and define

$$H_m^* = \sum_{i=1}^{G} (\tilde{\boldsymbol{x}}_i - \tilde{\boldsymbol{x}})' \boldsymbol{W}_i (\tilde{\boldsymbol{x}}_i - \tilde{\boldsymbol{x}})$$

where

$$\boldsymbol{W}_i = n_i \boldsymbol{S}_i^{-1}$$

$$\tilde{\boldsymbol{x}}_i = \frac{2}{n_i(n_i + 1)} \boldsymbol{x}_{in_i} + \frac{n_i - 1}{n_i(n_i + 1)} \bar{\boldsymbol{x}}_i$$

and

$$\tilde{\boldsymbol{x}} = [\sum_{i=1}^{G} \boldsymbol{W}_i]^{-1} \sum_{i=1}^{G} \boldsymbol{W}_i \tilde{\boldsymbol{x}}_i \quad .$$

The statistic $\underline{H}_m^*$ is approximately distributed as chi-square with p(G-1) degrees of freedom.

Invariance Property of the Test Statistics

Samples in this experiment were selected from either a contaminated population or an uncontaminated population. The subset of populations labeled uncontaminated had the identity matrix (**I**) as their common dispersion matrix. The subset of populations categorized as contaminated had a common diagonal matrix (**D**). That these matrix forms entail no loss of generality beyond the limited form of heteroscedasticity investigated is due to (a) a well known theorem by Anderson (1958) and (b) the invariance characteristic of the test statistics.

First, denote by $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ the pxp dispersion matrices for populations i and j, respectively. Since $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ are

positive definite, there exists a pxp nonsingular matrix $\mathbf{T}$ such that $\mathbf{T\Sigma_i T'} = \mathbf{I}$ and $\mathbf{T\Sigma_j T'} = \mathbf{D}$, where $\mathbf{I}$ is a pxp identity matrix and $\mathbf{D}$ is a pxp diagonal matrix (Anderson, 1958). Hence, when the design includes two population subsets with common dispersion matrices within a given subset, including only diagonal matrices in each simulated experiment is not an additional limitation on generalizability.

Second, the test statistics are invariant with respect to transformations where $\mathbf{T}$ is a pxp nonsingular transformation.

Brown-Forsythe Generalizations

Let $\mathbf{y}_{ij} = \mathbf{Tx}_{ij}$. Let $\overline{\mathbf{y}}_i, \mathbf{S}_i^*$ denote the sample mean vector and sample dispersion matrix for $\mathbf{y}_{ij}$ in the ith sample. It is well known that $\overline{\mathbf{y}}_i = \mathbf{T}\overline{\mathbf{x}}_i$ and $\mathbf{S}_i^* = \mathbf{TS_iT'}$. Let $\mathbf{H}^*$ and $\mathbf{M}^*$ be calculated by using $\overline{\mathbf{y}}_i$ and $\mathbf{S}_i^*$. It is well known that $\mathbf{H}^* = \mathbf{THT'}$. Now

$$\boldsymbol{M}^* = \sum_{i=1}^{G} (1 - \frac{n_i}{N}) \boldsymbol{S}_i^*$$

$$= \sum_{i=1}^{G} (1 - \frac{n_i}{N}) \boldsymbol{T} \boldsymbol{S}_i \boldsymbol{T}'$$

$$= \boldsymbol{T}[\sum_{i=1}^{G} (1 - \frac{n_i}{N}) \boldsymbol{S}_i] \boldsymbol{T}' = \boldsymbol{T} \boldsymbol{M} \boldsymbol{T}' .$$

For the modified Hotelling-Lawley trace criterion

$$U^* = trace\{ \boldsymbol{H}^* [\frac{f}{h} \boldsymbol{M}^*]^{-1} \}$$

$$= trace\{\boldsymbol{T}\,\boldsymbol{H}\,\boldsymbol{T}'\ [\ \frac{f}{h}\ \boldsymbol{T}\,\boldsymbol{M}\,\boldsymbol{T}'\ ]^{-1}\ \}$$

$$= trace\ \{\ \boldsymbol{H}[\ (\frac{f}{h})\,\boldsymbol{M}]^{-1}\ \}\quad .$$

Similarly, for the modified Pillai-Bartlett trace criterion

$$V^{*} = trace\{\ \boldsymbol{H}^{*}[\boldsymbol{H}^{*} + (\frac{f}{h})\,\boldsymbol{M}^{*}]^{-1}\}$$

$$= trace\{\boldsymbol{T}\,\boldsymbol{H}\,\boldsymbol{T}'[\boldsymbol{T}\,\boldsymbol{H}\,\boldsymbol{T}' + (\frac{f}{h})\,\boldsymbol{T}\,\boldsymbol{M}\,\boldsymbol{T}']^{-1}\}$$

$$= trace\{\boldsymbol{H}[\boldsymbol{H}+\frac{f}{h}\boldsymbol{M}]^{-1}\}\quad .$$

For the modified Wilks likelihood ratio criterion,

$$L^{*} = \frac{|\ \frac{f}{h}\ \boldsymbol{M}^{*}\ |}{|\ \boldsymbol{H}^{*} + \frac{f}{h}\ \boldsymbol{M}^{*}\ |}$$

$$= \frac{|\ \frac{f}{h}\ \boldsymbol{T}\,\boldsymbol{M}\,\boldsymbol{T}'\ |}{|\ \boldsymbol{T}\,\boldsymbol{H}\,\boldsymbol{T}' + \frac{f}{h}\ \boldsymbol{T}\,\boldsymbol{M}\,\boldsymbol{T}'\ |}$$

$$= \frac{|\boldsymbol{T}|\,|\frac{f}{h}\boldsymbol{M}|\,|\boldsymbol{T}'|}{|\boldsymbol{T}|\,|\boldsymbol{H} + \frac{f}{h}\boldsymbol{M}|\,|\boldsymbol{T}'|}$$

$$= \frac{|\frac{f}{h}\boldsymbol{M}|}{|\boldsymbol{H} + \frac{f}{h}\boldsymbol{M}|}\quad .$$

Wilcox Generalization

To show the invariance of $\underline{H}_m^{*}$, the following results are useful:

1. $W_i^* = n_i[S_i^*]^{-1} = [T']^{-1}n_i S_i^{-1} T^{-1}$.

2. $\sum_{i=1}^{G} W_i^* = (T')[\sum_{i=1}^{G} n_i S_i^{-1}]T^{-1} = (T')[\sum_{i=1}^{G} W_i]T^{-1}$.

3. $\tilde{y}_i = T\tilde{x}_i$.

4. $\tilde{y} = (\sum_{i=1}^{G} W_i^*)^{-1}(\sum_{i=1}^{G} W_i^*\tilde{y}_i)$,

$$= [(T')\{\sum_{i-1}^{G} W_i\}T^{-1}]^{-1}\sum_{i=1}^{G}(T')^{-1}n_i S_i^{-1}T^{-1}T\tilde{x}_i \quad ,$$

$$= T(\sum_{i=1}^{G} W_i)^{-1}(T')^{-1}T'\sum_{i=1}^{G} n_i S_i^{-1}\tilde{x}_i \quad ,$$

$$= T(\sum_{i=1}^{G} W_i)^{-1}\sum_{i=1}^{G} W_i\tilde{x}_i = T\tilde{x} \quad .$$

Using results 1-4, $\underline{H}_m^*$ is shown to be invariant as follows:

$$H_m^* = \sum_{i=1}^{G}(T\tilde{x}_i - T\tilde{x})'W^*{}_i(T\tilde{x}_i - T\tilde{x}) \quad ,$$

$$= \sum_{i=1}^{G}[T(\tilde{x}_i - \tilde{x})]'(T')^{-1}W_{i_i}T^{-1}[T(\tilde{x}_i - \tilde{x})] \quad ,$$

$$= \sum_{i=1}^{G}(\tilde{x}_i - \tilde{x})'W_i(\tilde{x}_i - \tilde{x}) \quad .$$

Therefore, the test statistics $\underline{U}^*$, $\underline{V}^*$, $\underline{L}^*$ and $\underline{H}_m^*$ are invariant to nonsingular transformation. Thus, there is no

loss of generality by solely using diagonal matrices to simulate experiments in which there are only two sets of dispersion matrices. It should be noted, however, when there are more than two sets of differing dispersion matrices, the matrices cannot always be simultaneously diagonalized by a transformation matrix **T**.

Design

Eight factors were considered in the study. These are described in the following paragraphs.

Distribution type (DT). Two types of distributions--normal and exponential--were included in the study. Pearson and Please (1975) suggested that studies of robustness should focus on distributions with skewness and kurtosis having magnitudes less than 0.8 and 0.6, respectively. However, there is evidence to suggest these boundaries are unnecessarily restrictive. For example, Kendall and Stuart (1963, p.57) reported the age at time of marriage for over 300,000 Australians. The skewness and kurtosis were 2.0 and 8.3, respectively. Micceri (1989) investigated the distributional characteristics of 440 achievement and psychometric measures. Of these 440 data sets, 15.2% had both tails with weights at or about Gaussian, 49.1% had at least one extremely heavy tail, and 18.0% had both tail weights less than Gaussian. The Micceri study found 28.4% of the distributions were relatively symmetric, 40.7% were classified as being moderately asymmetric, and 30.7% were classified as

being extremely asymmetric. Of the distributions considered, 11.4% were classified within a category having skewness as extreme as 2.00. The Micceri study underscores the common occurrence of distributions that are non-normal. Further, the Micceri study suggests the Pearson and Please criterion may be too restrictive.

For the normal distribution the coefficients of skewness $(\mu_3^2/\mu_2^3)^{1/2}$ and kurtosis $(\mu_4/\mu_2^2 - 3)$ are respectively 0.00 and 0.00. For the exponential distribution the coefficients of skewness and kurtosis are respectively 2.00 and 6.00. The Micceri study provides evidence that the proposed normal and exponential distributions are reasonable representations of data that may be found in educational research.

Number of dependent variables (p). Data were generated to simulate experiments in which there are p=3 or p=6 dependent variables. This choice is reasonably consistent with the range of variables commonly examined in educational research (Algina & Oshima, 1990; Algina & Tang, 1988; Hakstian, Roed, & Lind, 1979; Lin, 1991; Olson, 1974; Tang, 1989).

Number of populations sampled (G). Data were generated to simulate experiments in which there is sampling from either G=3 or G=6 populations. Dijkstra & Werter (1981) simulated experiments with G equal to 3, 4, and 6. Olson (1974) simulated experiments with G equal to 2, 3, 6, and 10. Multivariate experiments with a large number of groups seem to

be rare in educational research (Tang, 1989). Hence, the chosen number of populations sampled should provide reasonably adequate examination of this factor.

Degree of the sample size ratio (NR). Only unequal sample sizes are used in the study. Sample size ratios were chosen to range from small to moderately large. The basic ratios of $n_1:n_2:n_3$ used in the simulation when sampling from three different populations are given in Table 2. Similarly, the ratios of $n_1:...:n_6$ used in the simulation when sampling from six different populations are given in Table 3. Fairly large ratios were used in Algina and Tang's (1988) study, with an extreme ratio of 5:1. In experimental and field studies, it is common to have sample-size ratios between 1:1 and 2:1 (Lin, 1991). Olson (1974) examined only the case of equal-sized samples.

Since error rates increase as the degree of the sample size ratio increases (Algina & Oshima, 1990), if nominal error rates are excessively exceeded using small to moderately large sample size ratios, then the procedure presumably will have difficulty with extreme sample size ratios. Conversely, if the procedure performs well under this range of sample size ratios, then it should work well for equal sample size ratios and the question of extreme sample size ratios is still open. Hence, sample size ratios were chosen under the constraint $n_{[G]}:n_{[1]}$ is less extreme than 2:1, where $n_{[1]}$ is the smallest

Table 2

Sample Size Ratios (n_1:n_2:n_3)

n_1 :	n_2 :	n_3
1	1	1.3
1	1	2
1	1.3	1.3
1	2	2

Table 3

Sample Size Ratios (n_1:...:n_6)

n_1 :	n_2 :	n_3 :	n_4 :	n_5 :	n_6
1	1	1	1	1.3	1.3
1	1	1	1	2	2
1	1	1.3	1.3	1.3	1.3
1	1	2	2	2	2

sample size and $n_{[G]}$ is the largest sample size of the G populations sampled. In some cases these basic ratios could not be maintained because of the restriction of the ratio of total sample size to number of dependent variables. Departure from these basic ratios was minimized.

Form of the sample size ratio (NRF). When there are three groups, either the sample size ratio is of the form $n_1 = n_2 < n_3$ and is denoted by NRF=1 or the sample size ratio is of the form $n_1 < n_2 = n_3$ and is denoted by NRF=2. When there are six groups, either the sample size ratio is of the form $n_1 = n_2 = n_3 = n_4 < n_5 = n_6$ and is denoted by NRF=1 or the sample size ratio is of the form $n_1 = n_2 < n_3 = n_4 = n_5 = n_6$ and is denoted NRF=2.

Ratio of total sample size to number of dependent variables (N:p). The ratios chosen were N:p=10 and N:p=20. Hakstian, Roed, and Lind (1979) simulated experiments with N:p equal to 6 or 20. With some notable exceptions (Algina & Tang, 1988; Lin, 1991) current studies tend to avoid N:p smaller than 10. Yao's test (which is generally more robust than James's first-order test) should have N:p at least 10 to be robust (Algina & Tang, 1988). With G>2, Lin (1991) reasoned it seems likely that N:p will need to be at least 10 for robustness to be obtained. An upper limit of 20 was chosen to represent moderately large experiments. These selections result in a minimum total sample size of N=30 and a maximum total sample size of N=120.

Degree of heteroscedasticity (d). Each population with dispersion matrix equal to a pxp identity matrix (**I**) will be called an uncontaminated population. Each population with a pxp diagonal dispersion matrix (**D**) with at least one diagonal element not equal to one will be called a contaminated population. The forms of the dispersion matrices, which depend upon the number of dependent variables, are shown in Table 4. Two levels of d, $d=\sqrt{2}$ and d=3.0, were used to simulate the degree of heteroscedasticity of the dispersion matrices. Olson (1974) simulated experiments with d equal to 2.0, 3.0, and 6.0. Algina and Tang (1988) simulated experiments with d equal to 1.5, 2.0, 2.5, and 3.0. Tang (1989) chose d equal to $\sqrt{1.5}$ and 3.0. Algina and Oshima (1990) selected d equal to 1.5 and 3.0. For this study, $d=\sqrt{2}$ was used to simulate a small degree of heteroscedasticity and d=3.0 was selected to represent a larger degree of heteroscedasticity. These values were selected to represent a range of heteroscedasticity more likely to be common in educational experiments (Tang, 1989).

Relationship of sample size to dispersion matrices (S). Both positive and negative relationships between sample size and dispersion matrices were investigated. In the positive relationship, the larger samples correspond to **D**. In the negative relationship, the smaller samples correspond to **D**. These relationships for G=3 and G=6 are summarized in Table 5 and Table 6, respectively.

Table 4

Forms of Dispersion Matrices

Matrix	p=3	p=6
D	Diag{ $1,d^2,d^2$ }	Diag{ $1,1,d^2,d^2,d^2,d^2$ }
I	Diag{1,1,1}	Diag{1,1,1,1,1,1}

Table 5

Relationship of Sample Size to Heteroscedasticity (G=3)

Sample Size Ratios			Relationship	
n_1 :	n_2 :	n_3	Positive	Negative
1	1	1.3	**IID**	**DDI**
1	1	2	**IID**	**DDI**
1	1.3	1.3	**IDD**	**DII**
1	2	2	**IDD**	**DII**

Table 6

Relationship of Sample Size to Heteroscedasticity (G=6)

Sample Size Ratios						Relationship	
n_1 :	n_2 :	n_3 :	n_4 :	n_5 :	n_6	Positive	Negative
1	1	1	1	1.3	1.3	**IIIIDD**	**DDDDII**
1	1	1	1	2	2	**IIIIDD**	**DDDDII**
1	1	1.3	1.3	1.3	1.3	**IIDDDD**	**DDIIII**
1	1	2	2	2	2	**IIDDDD**	**DDIIII**

Design Layout. The sample sizes were determined once values of p, G, N:p, NRF, and NR were specified. These sample sizes are summarized for G=3 and G=6 in Table 7 and Table 8, respectively. Each of these 32 conditions were crossed with two distributions, two levels of heteroscedasticity, and two relationships of sample size to dispersion matrices to generate 256 experimental conditions from which to draw conclusions regarding the competitiveness of the proposed statistics to the established Johansen procedure.

Simulation Procedure

The simulation was conducted as 256 separate runs, one for each condition, with 2000 replications per condition. For each condition, the performance of Johansen's test ($\underline{J}$), the two variations of the modified Hotelling-Lawley test ($\underline{U}_1^*$, $\underline{U}_2^*$), the modified Pillai-Bartlett test ($\underline{V}^*$), the modified Wilks test ($\underline{L}^*$), and the modified Wilcox test ($\underline{H}_m^*$) were evaluated using the generated data.

For the ith sample, an $n_i$xp (i=1,...,G) matrix of uncorrelated pseudo-random observations was generated (using PROC IML in SAS) from the target distribution--normal or exponential. When the target distribution was an exponential, the random observations on each of the p variates were standardized using the population expected value and standard deviation. Hence, within each uncontaminated population, all

Table 7

Sample Sizes (G=3)

p	G	N:p	N	n_1	n_2	n_3
3	3	10	30	9	9	12
				7	7	16
				8	11	11
				6	12	12
		20	60	18	18	24
				15	15	30
				16	22	22
				12	24	24
6	3	10	60	18	18	24
				15	15	30
				16	22	22
				12	24	24
		20	120	36	36	48
				30	30	60
				32	44	44
				24	48	48

Note. N is occasionally altered to maintain the ratio as closely as manageable.

Table 8

Sample Sizes (G=6)

p	G	N:p	N	n_1	n_2	n_3	n_4	n_5	n_6
3	6	10	30	5	5	5	5	6	6
				4	4	4	4	8	8
				4	4	5	5	5	5
				4	4	6	6	6	6
		20	60	9	9	9	9	12	12
				7	7	7	7	16	16
				8	8	11	11	11	11
				6	6	12	12	12	12
6	6	10	60	9	9	9	9	12	12
				7	7	7	7	16	16
				8	8	11	11	11	11
				7	7	12	12	12	12
		20	120	18	18	18	18	24	24
				15	15	15	15	30	30
				16	16	22	22	22	22
				12	12	24	24	24	24

Note. N is occasionally altered to maintain the ratio as closely as manageable.

the p variates were identically distributed with mean equal to zero, variance equal to one, and all covariances among the p variates equal to zero.

Each $n_i$xp matrix of observations corresponding to a contaminated population was post multiplied by an appropriate **D** to simulate dispersion heteroscedasticity.

For each replication, the data were analyzed using Johansen's test, the two variations of the modified Hotelling-Lawley trace criterion, the modified Pillai-Bartlett trace criterion, the modified Wilks likelihood ratio criterion, and the modified Wilcox test. The proportion of 2000 replications that yielded significant results at $\alpha = 0.05$ were recorded.

Summary

Two distribution types [DT=normal or exponential], two levels of dependent variables (p=3 or 6), two levels of populations sampled (G=3 or 6), two levels of the form of the sample size ratio, two levels of the degree of the sample size ratio, two levels of ratio of total sample size to number of dependent variables (N:p=10 or 20), two levels of degree of heteroscedasticity (d=$\sqrt{2}$ or 3.0), and two levels of the relationship of sample size to dispersion matrices (S=positive or negative condition) combine to give 256 experimental conditions. The Johansen test ($\underline{J}$), the two variations of the modified Hotelling-Lawley test ($\underline{U}_1^*$, $\underline{U}_2^*$), the modified Pillai-Bartlett test ($\underline{V}^*$), the modified Wilks test ($\underline{L}^*$), and the

modified Wilcox test ($\underline{H}_m^*$) were applied to each of these experimental conditions. Generalizations of the behavior of these tests will be based upon the collective results of these 256 experimental conditions.

CHAPTER 4
RESULTS AND DISCUSSION

In this chapter analyses of $\hat{\tau}$ for α=.05 are presented. Results with regard to $\hat{\tau}$ for α=.01 and for α=.10 are similar. The analyses are based on data presented in the Appendix.

Distributions of $\hat{\tau}$ for the six tests are depicted in Figures 1 to 6. In each of these six figures, the interval labelled .05 denotes $.0250 \leq \hat{\tau} \leq .0749$, the interval labelled .10 denotes $.0750 \leq \hat{\tau} \leq .1249$, and so forth. From these figures it is clear that in terms of controlling Type I error rates (a) the performance of the Johansen ($\underline{J}$) and modified Wilcox ($\underline{H}_m^*$) tests are similar; (b) the performance of the first modified Hotelling-Lawley ($\underline{U}_1^*$), second modified Hotelling-Lawley ($\underline{U}_2^*$), modified Pillai-Bartlett ($\underline{V}^*$), and modified Wilks ($\underline{L}^*$) tests are similar; (c) the performance of these two sets of tests greatly differ from one another; (d) the performance of the Johansen test is superior to that of the Wilcox generalization; and (e) the performance of each of the Brown-Forsythe generalizations is superior to that of either the Johansen test or Wilcox generalization. Because the performance of the Johansen and modified Wilcox tests were so different from that of the Brown-Forsythe generalizations, separate analyses were conducted for each of these two sets of tests. For each separate set of tests, analysis of variance

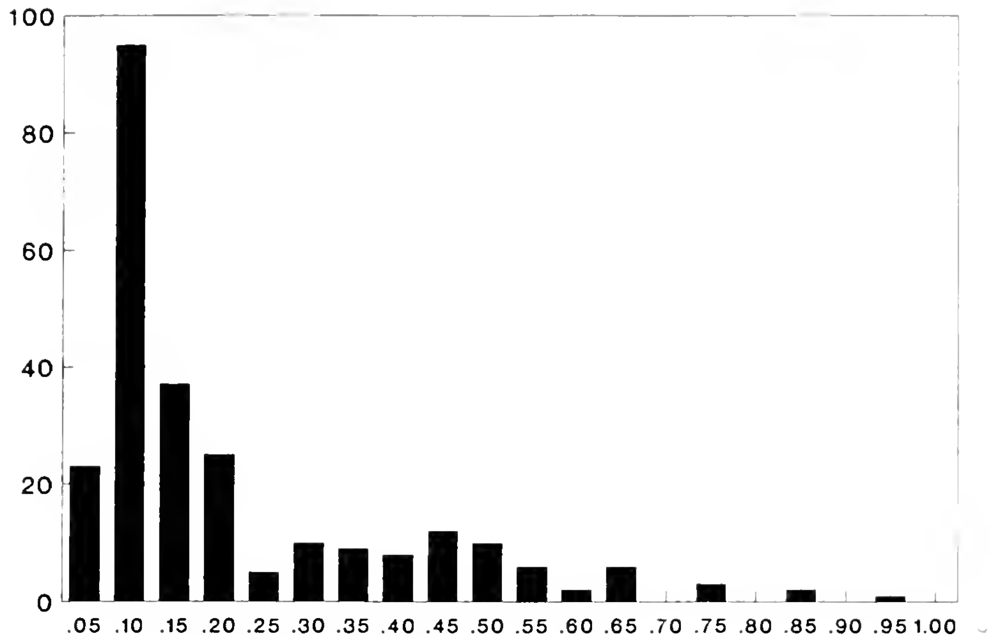

Figure 1

<u>Frequency Histogram of Estimated Type I Error Rates for the Johansen Test</u>

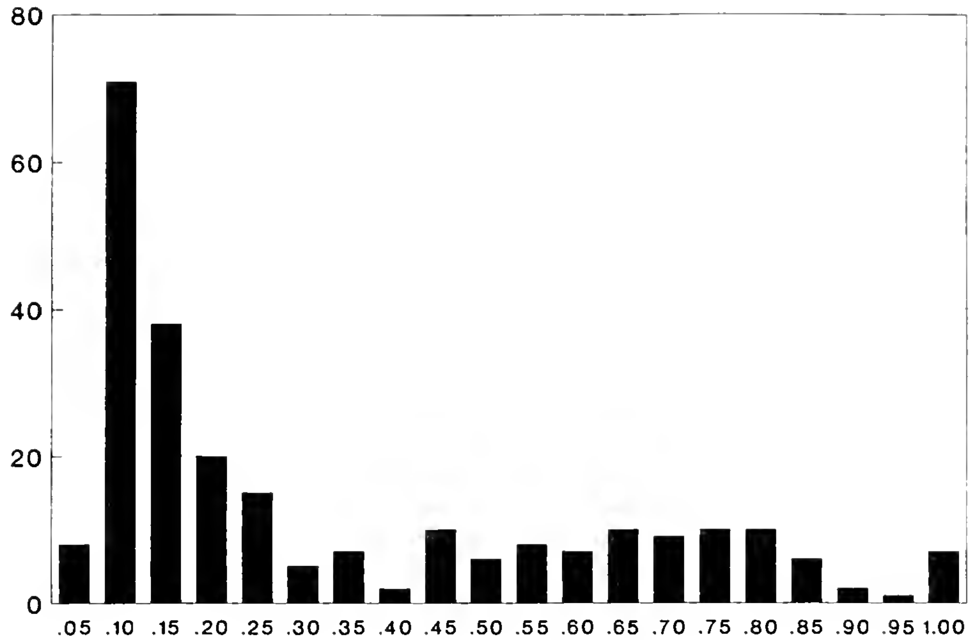

Figure 2

<u>Frequency Histogram of Estimated Type I Error Rates for the Modified Wilcox Test</u>

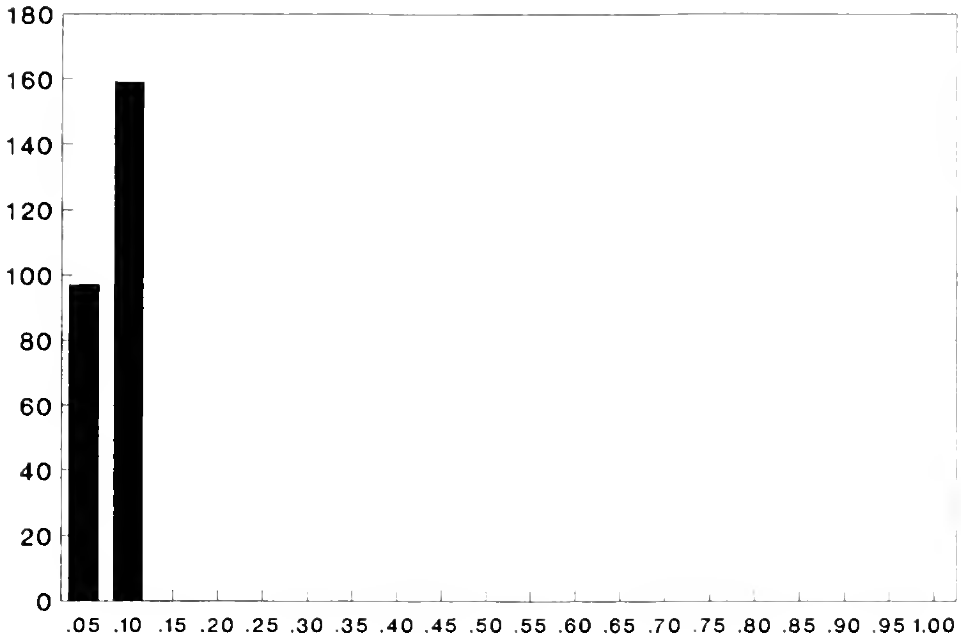

Figure 3

Frequency Histogram of Estimated Type I Error Rates for the First Modified Hotelling-Lawley Test

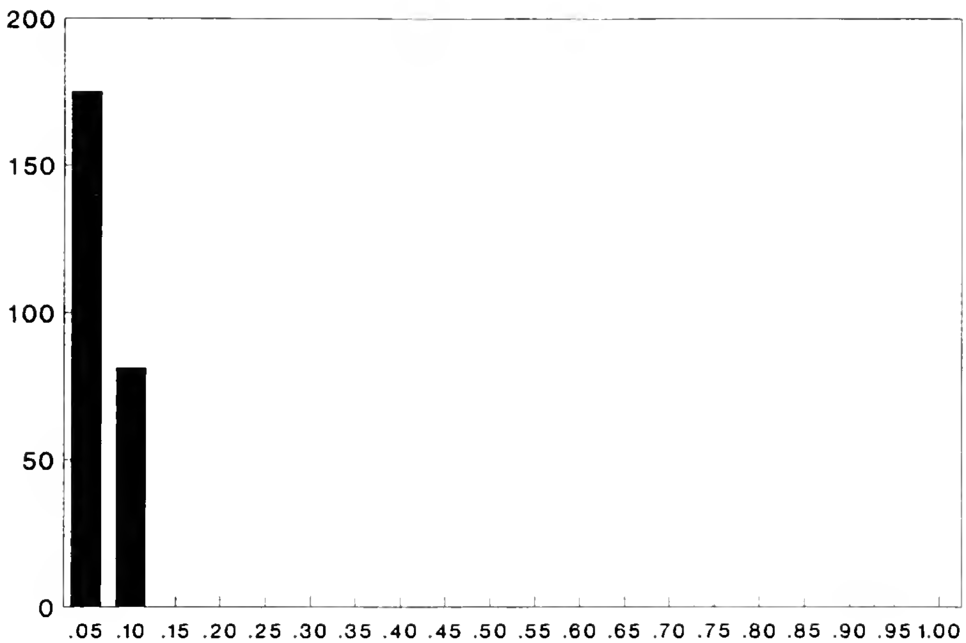

Figure 4

<u>Frequency Histogram of Estimated Type I Error Rates for the Second Modified Hotelling-Lawley Test</u>

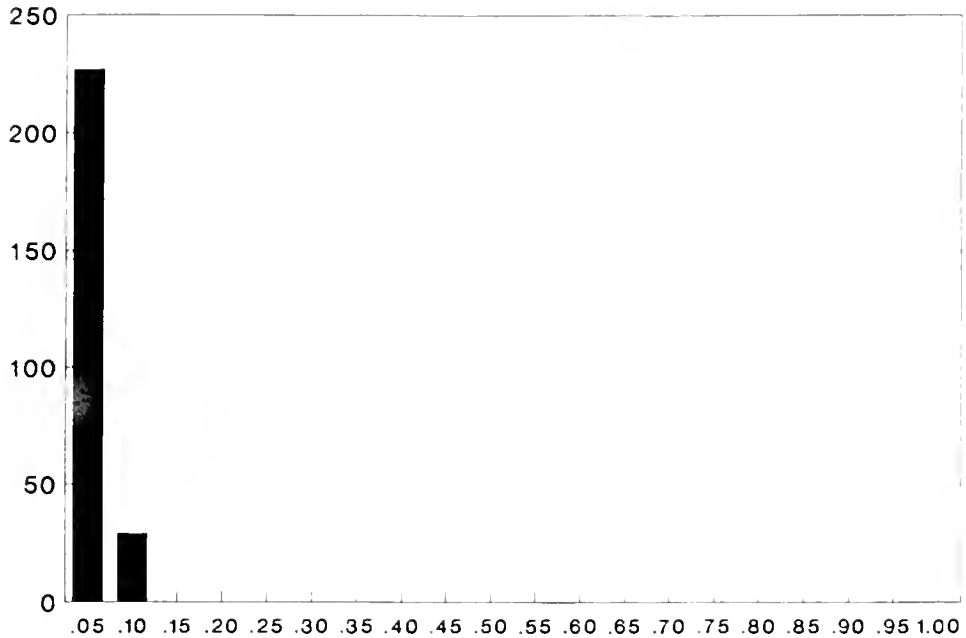

Figure 5

Frequency Histogram of Estimated Type I Error Rates for the Modified Pillai-Bartlett Test

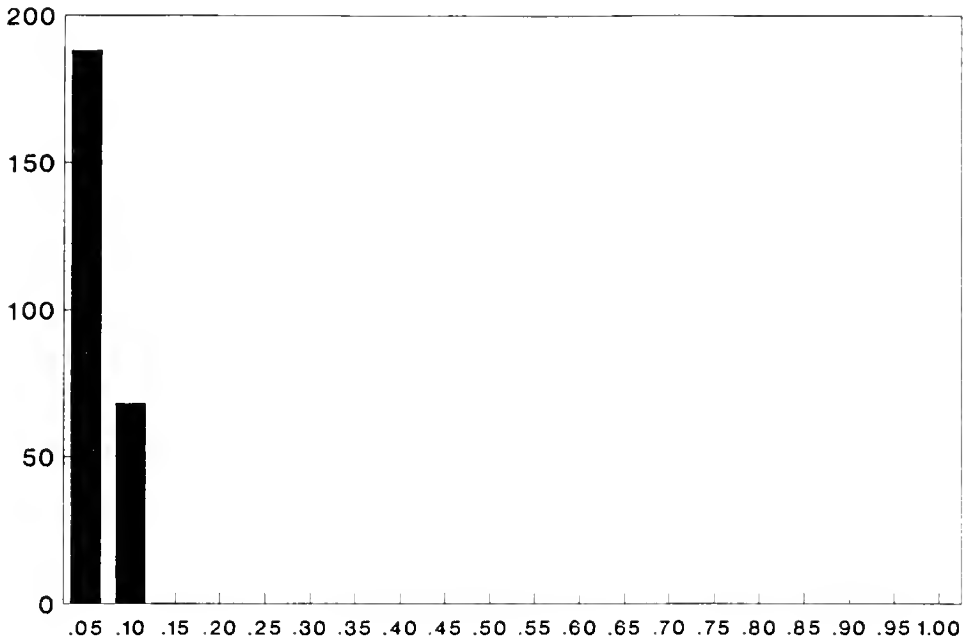

Figure 6

Frequency Histogram of Estimated Type I Error Rates for the Modified Wilks Test

was used to investigate the effect on $\hat{\tau}$ of the following factors: Distribution Type (DT), Number of Dependent Variables (p), Number of Populations Sampled (G), Degree of the Sample Size Ratio (NR), Form of the Sample Size Ratio (NRF), Ratio of Total Sample Size to Number of Dependent Variables (N:p), Degree of Heteroscedasticity (d), Relationship of Sample Size to Dispersion Matrices (S), and Test Criteria (T).

Brown-Forsythe Generalizations

Because there are nine factors, initial analyses were conducted to determine which effects to enter into the analysis of variance model. A forward selection approach was used, with all main effects entered first, followed by all two-way interactions, all three-way interactions, and all four-way interactions. Because R^2 was .96 for the model with four-way interactions, more complex models were not examined. The R^2 for all models are shown in Table 9. The model with main effects and two-way through four-way interactions was selected.

Variance components were computed for each main effect, two-way, three-way, and four-way interaction. The variance component (θ_i, i=1,...,255) for each effect was computed using the formula 10^6(MSEF-MSE)/(2^x), where MSEF was the mean square for that given effect, MSE was the mean square error for the four-factor interaction model, and 2^x was the number of levels for the factors not included in that given effect. Negative

Table 9

Magnitudes of R^2 for Main Effects, Two-Way Interaction, Three-Way Interaction, and Four-Way Interaction Models when using the Four Brown-Forsythe Generalizations

Highest-Order Terms	R^2
Main Effects	0.52
Two-Way Interactions	0.77
Three-Way Interactions	0.89
Four-Way Interactions	0.96

variance components were set to zero. Using the sum of these variance components plus MSEx10^6 as a measure of total variance, the proportion of total variance in estimated Type I error rates was computed for the ith effect (i=1,...,255) using the formula $\theta_i/[(\theta_1+...+\theta_{255}) + 10^6\text{MSE}]$. Shown in Table 10 are effects that (a) were statistically significant and (b) accounted for at least 1% of the total variance in estimated Type I error rates.

Because N:p, T, G, and GxT are among the largest effects and--in contrast to factors such as d and DT--do not have to be inferred from data, their effects were examined by calculating percentiles of $\hat{\tau}$ for each combination of G and N:p. These percentiles should provide insight into the functioning of the four tests. The DTxNRFxSxd interaction was significant and the second largest effect. Consequently the effects of the four factors involved in this interaction were examined by constructing cell mean plots involving all combinations of the four factors. Other interaction effects with large variance components that included these factors were checked and did not change the findings significantly. The DTxG interaction will be examined because it (a) accounts for 4.0% of the total variance in estimated Type I error rates and (b) is not explained in terms of either the effect of T, N:p, and G or the effect of DT, NRF, S, and d. The factor p has neither a large main effect or large interactions with any other factors. However, because it accounted for 1.5% of the

Table 10

Variance Components for the First Modified Hotelling-Lawley, Second Modified Hotelling-Lawley, Modified Pillai-Bartlett, and Modified Wilks Tests

Effect	θ	Percent of Variance
N:p	81	8.3
DTxNRFxSxd	76	7.8
T	66	6.8
DTxNRFxS	55	5.7
NRFxSxd	46	4.8
DTxd	41	4.2
DTxG	39	4.0
NRFxS	35	3.6
G	33	3.4
GxT	31	3.2
DTxGxd	25	2.6
DTxGxNRFxS	25	2.5
Sxd	23	2.4
d	22	2.3
S	21	2.2
NRFxSxdxT	18	1.8
p	15	1.5

Table 10--continued.

Effect	θ	Percent of Variance
pxNRFxSxd	13	1.3
DTxGxN:pxd	13	1.3
GxN:p	13	1.3
NRFxSxT	11	1.2
DTxNRxS	10	1.0
dxT	10	1.0
DTxS	10	1.0
All Others	<10	<1.0

variance, its effect was examined by inspecting cell means for p=3 and p=6. Finally, the influence of the degree of the sample size ratios (NR) was minimal. The NR main effect accounted for only .1% of the total variance in estimated Type I error rates. The three-way interaction DTxNRxS was the effect with the largest variance component which included NR and it still only accounted for 1.0% of the total variance in estimated Type I error rates.

Effect of T, N:p, and G. Percentiles for $\underline{U}_1^*$ and $\underline{U}_2^*$ are displayed in Table 11; percentiles for $\underline{V}^*$ and $\underline{L}^*$ are shown in Table 12. Using Bradley's liberal criterion ($.5\alpha \le \hat{\tau} \le 1.5\alpha$), the following patterns emerge regarding control of Type I error rates for the Brown-Forsythe generalizations: (a) the first modified Hotelling-Lawley test ($\underline{U}_1^*$) was adequate when N:p was 10; however, the test tended to be liberal when N:p was 20; (b) the second modified Hotelling-Lawley test ($\underline{U}_2^*$) was adequate when either N:p was 10 and G was 3 or when N:p was 20 and G was 6; (c) the second modified Hotelling-Lawley test tended to be conservative when N:p was 10 and G was 6 whereas the test tended to be slightly liberal when N:p was 20 and G was 3; (d) the modified Pillai-Bartlett test ($\underline{V}^*$) was adequate when N:p was 20 and G was 3; (e) the modified Pillai-Bartlett test tended to be conservative when N:p was 10 or when N:p was 20 and G was 6; (f) the modified Wilks test was adequate when N:P was 10 and G was 3 or when N:p was 20; and (g) the

Table 11

Percentiles of $\hat{\tau}$ for the First Modified Hotelling-Lawley Test (U_1^*) and Second Modified Hotelling-Lawley Test (U_2^*) for Combinations of Ratio of Total Sample Size to Number of Dependent Variables (N:p) and Number of Populations Sampled (G)

		G (N:p=10)		G (N:p=20)	
Test	Percentile	3	6	3	6
U_1^*	95th	.0795*	.0770	.0855*	.0885*
	90th	.0710	.0715	.0795*	.0835*
	75th	.0555	.0595	.0610	.0708
	50th	.0505	.0500	.0538	.0625
	25th	.0430	.0398	.0493	.0540
	10th	.0375	.0315	.0460	.0490
	5th	.0345	.0295	.0435	.0485
U_2^*	95th	.0730	.0510	.0815*	.0710
	90th	.0625	.0460	.0785*	.0650
	75th	.0513	.0388	.0590	.0565
	50th	.0453	.0290	.0510	.0483
	25th	.0385	.0198*	.0470	.0388
	10th	.0325	.0140*	.0430	.0355
	5th	.0290	.0135*	.0405	.0330

Note. Percentiles denoted by an asterisk fall outside the interval [.5α,1.5α].

Table 12

Percentiles of $\hat{\tau}$ for the Modified Pillai-Bartlett Test (V^*) and Modified Wilks Test (L^*) for Combinations of Ratio of Total Sample Size to Number of Dependent Variables (N:p) and Number of Populations Sampled (G)

		G (N:p=10)		G (N:p=20)	
Test	Percentile	3	6	3	6
V^*	95th	.0555	.0365	.0695	.0510
	90th	.0495	.0310	.0660	.0500
	75th	.0430	.0258	.0533	.0455
	50th	.0370	.0210*	.0480	.0380
	25th	.0318	.0145*	.0425	.0315
	10th	.0240*	.0110*	.0365	.0275
	5th	.0200*	.0070*	.0345	.0235*
L^*	95th	.0705	.0465	.0780*	.0615
	90th	.0635	.0425	.0745	.0580
	75th	.0483	.0360	.0575	.0533
	50th	.0440	.0288	.0513	.0450
	25th	.0388	.0215*	.0455	.0375
	10th	.0330	.0155*	.0415	.0345
	5th	.0310	.0130*	.0405	.0325

Note. Percentiles denoted by an asterisk fall outside the interval [.5α, 1.5α].

modified Wilks test was conservative when N:p was 10 and G was 6.

Effect of DT, NRF, S, and d. As shown in Figure 7 and Figure 8, when data were sampled from a normal distribution, regardless of the form of the sample size ratio, mean $\hat{\tau}$ increased as degree of heteroscedasticity increased in the positive condition whereas mean $\hat{\tau}$ decreased as degree of heteroscedasticity increased in the negative condition. However, as shown in Figures 9 and 10, when data were sampled from an exponential distribution, mean $\hat{\tau}$ increased as degree of heteroscedasticity increased regardless of the relationship of sample sizes and dispersion matrices. The mean difference in $\hat{\tau}$ between the higher and lower degree of heteroscedasticity was greater in the positive condition when the sample was selected as in the first form of the sample size ratios whereas when the sample was selected as in the second form of the sample size ratio, the mean difference was greater in the negative condition. With data sampled from an exponential distribution the Brown-Forsythe generalizations tend to be conservative when (a) there was a slight degree of heteroscedasticity (that is, $d=\sqrt{2}$), (b) the degree of heteroscedasticity increased (d=3) and the first form of the sample size ratio was paired with the negative condition, or (c) the degree of heteroscedasticity increased and the second form of the sample size ratio was paired with the positive condition. With data sampled from an exponential

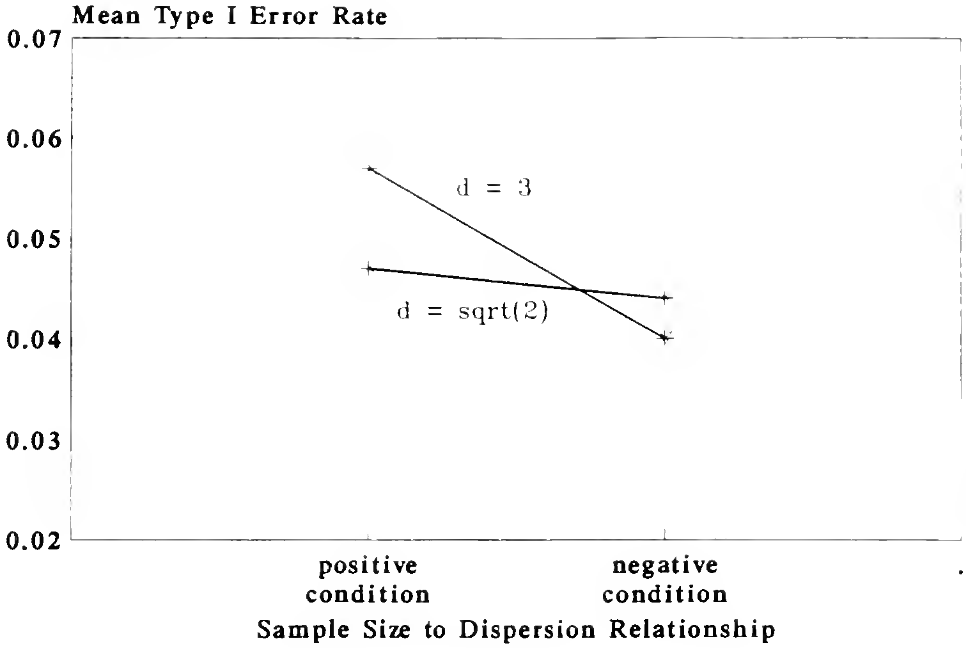

Figure 7

Estimated Type I Error Rates for the Two Levels of the Degree of Heteroscedasticity (d = $\sqrt{2}$ or 3) and Relationship of Sample Size to Dispersion Matrices (S = positive or negative condition) When Data Were Sampled as in the First Form of the Sample Size Ratio from an Normal Distribution

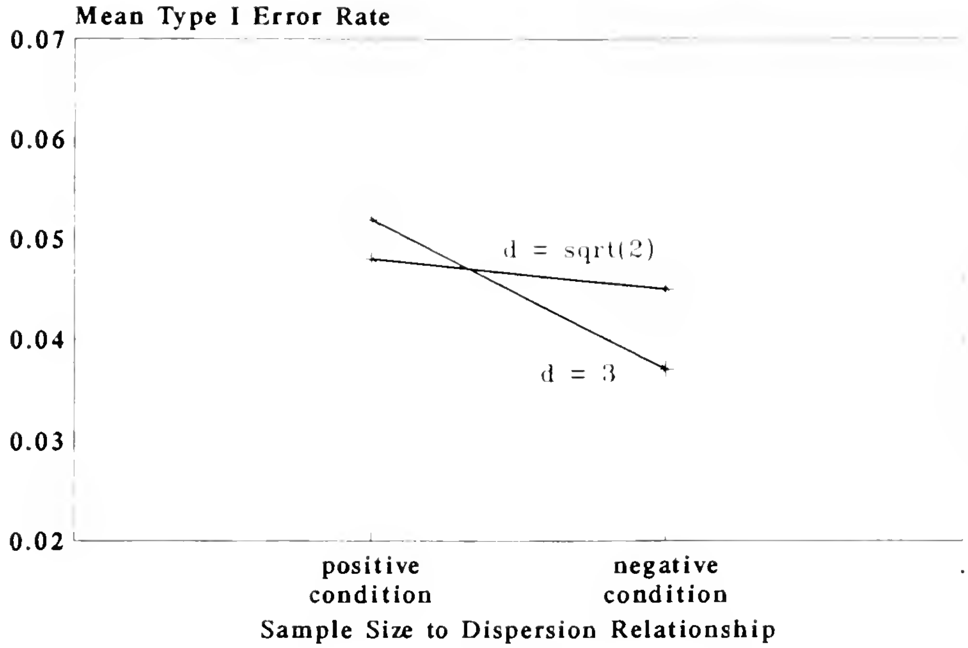

Figure 8

<u>Estimated Type I Error Rates for the Two Levels of the Degree of Heteroscedasticity (d = √2 or 3) and Relationship of Sample Size to Dispersion Matrices (S = positive or negative condition) When Data Were Sampled as in the Second Form of the Sample Size Ratio from an Normal Distribution</u>

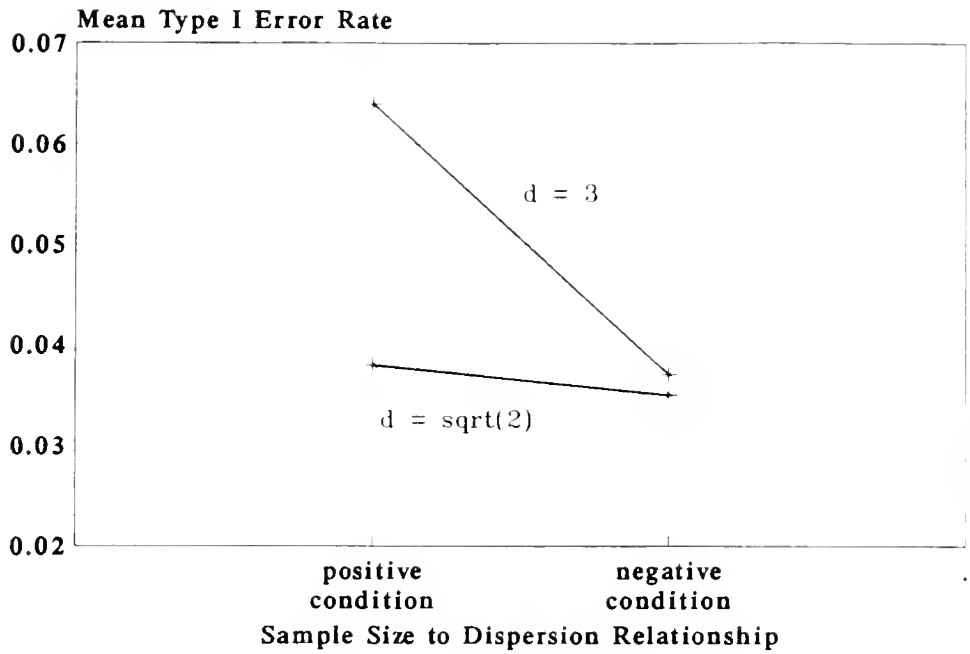

Figure 9

Estimated Mean Type I Error Rates for the Two Levels of the Degree of Heteroscedasticity (d = √2 or 3) and Relationship of Sample Size to Dispersion Matrices (S = positive or negative condition) When Data Were Sampled as in the First Form of the Sample Size Ratio from a Exponential Distribution

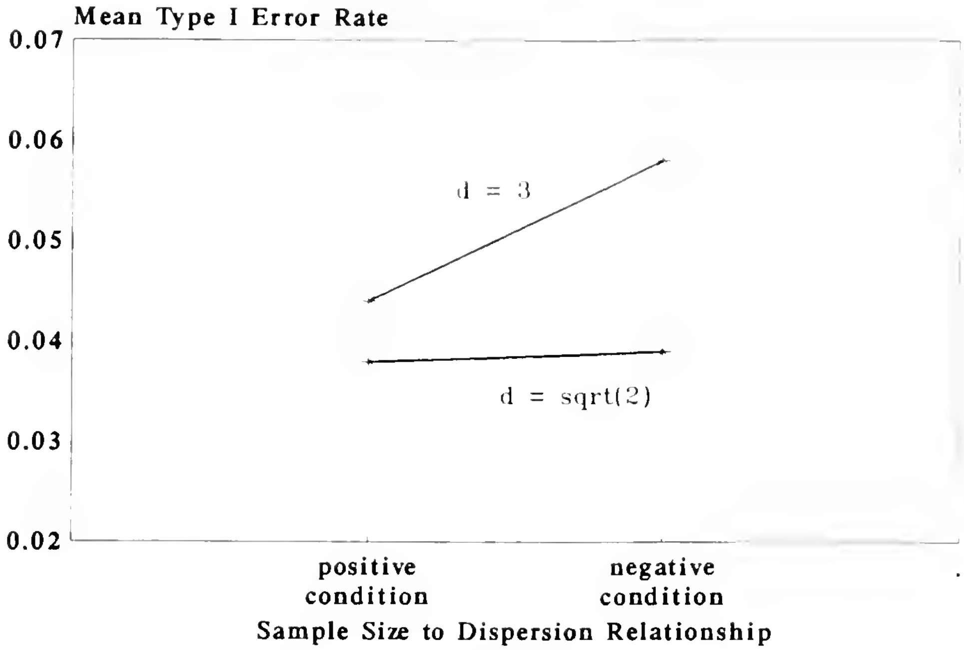

Figure 10

Estimated Mean Type I Error Rates for the Two Levels of the Degree of Heteroscedasticity (d = $\sqrt{2}$ or 3) and Relationship of Sample Size to Dispersion Matrices (S = positive or negative condition) When Data are Sampled as in the Second Form of the Sample Size Ratio from a Exponential Distribution

distribution, the Brown-Forsythe generalizations tended to be liberal when (a) the first form of the sample size ratio was paired with the positive condition, or (b) the second form of the sample size ratio was paired with the negative condition.

Effect of DTxG interaction. As shown in Figure 11, mean $\hat{\tau}$ for the Brown-Forsythe generalizations was nearer α when G was 3 than when G was 6, regardless of the type of distribution from which the data were sampled. When data were sampled from a normal distribution, the tests tended to be slightly conservative. Mean $\hat{\tau}$ was near α when data were sampled from an exponential distribution and G was 3. However, when data were sampled from an exponential distribution and G was 6, the Brown-Forsythe generalizations tended to be conservative.

Effect of p. Shown in Figure 12, mean $\hat{\tau}$ was near α for the Brown-Forsythe generalizations when p was 6. When p was 3, the tests tended to be slightly conservative.

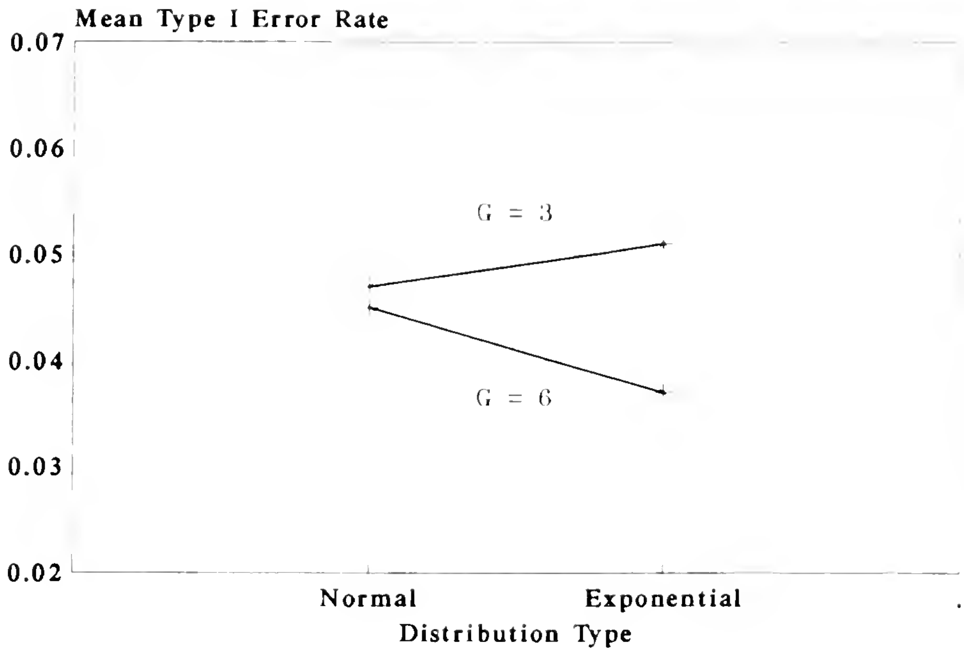

Figure 11

<u>Estimated Mean Type I Error Rates for Combinations of the Distribution Type and Number of Populations Sampled</u>

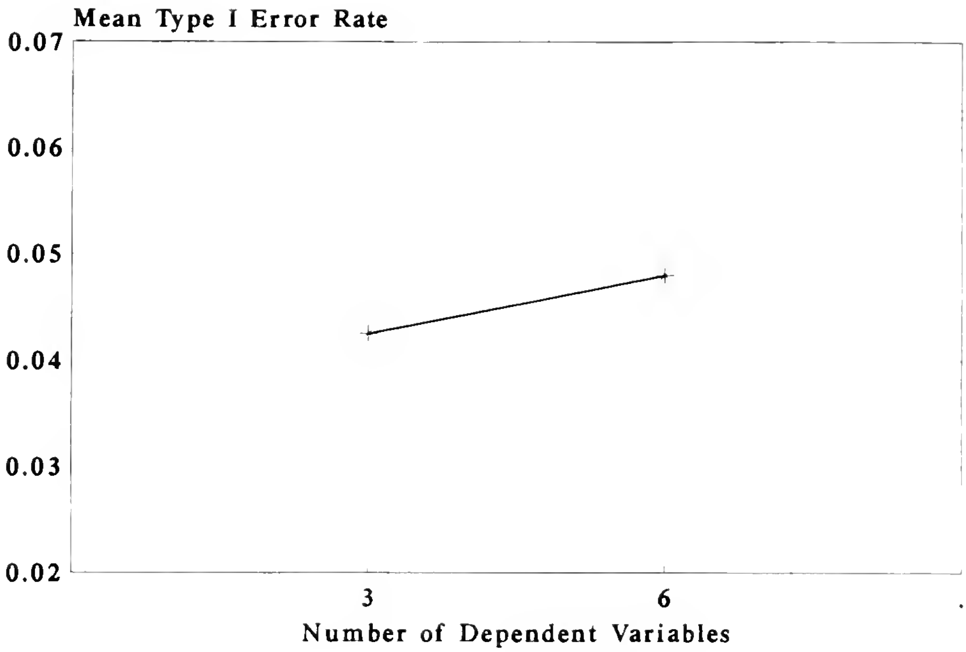

Figure 12

Estimated Mean Type I Error Rates for the Brown-Forsythe Generalizations for the Two Levels of the Number of Dependent Variables

Johansen Test and Wilcox Generalization

Because there are nine factors, initial analyses were conducted to determine which effects to enter into the analysis of variance model. A forward selection approach was used, with all main effects entered first, followed by all two-way interactions, all three-way interactions, and all four-way interactions. Because R^2 was .997 for the model with four-way interactions, more complex models were not examined. The R^2 for all models are shown in Table 13. The model with main effects and two-way through four-way interactions was selected.

Variance components were computed for each main effect, two-way, three-way, and four-way interaction. The variance component (θ_i, i=1,...,255) for each effect was computed using the formula $10^4(\text{MSEF}-\text{MSE})/(2^x)$, where MSEF was the mean square for that given effect, MSE was the mean square error for the four-factor interaction model, and 2^x was the number of levels for the factors not included in that given effect. Negative variance components were set to zero. Using the sum of these variance components plus MSEx10^4 as a measure of total variance, the proportion of total variance in estimated Type I error rates was computed for the ith effect (i=1,...,255) using the formula $\theta_i/[(\theta_1+\ldots+\theta_{255}) + 10^4\text{MSE}]$. Shown in Table 14 are effects that (a) were statistically significant and (b) accounted for at least 1% of the total variance in estimated Type I error rates.

Table 13

Magnitudes of R^2 for Main Effects, Two-Way Interaction, Three-Way Interaction, and Four-Way Interaction Models when using the Johansen Test and Wilcox Generalization

Highest-Order Terms	R^2
Main Effects	0.767
Two-Way Interactions	0.963
Three-Way Interactions	0.988
Four-Way Interactions	0.997

Table 14

Variance Components for the Johansen and Modified Wilcox Tests

Effect	θ	Percent of Variance
G	557	36.2
GxN:p	227	14.8
N:p	174	11.4
GxT	83	5.4
T	73	4.8
GxNRFxNR	34	2.2
p	26	1.7
pxG	26	1.7
GxNR	20	1.3
DT	18	1.2
NR	18	1.1
NRFxNR	17	1.1
All Others	<15	1.0

Because N:p, T, G, GxN:p, and GxT are among the largest effects and--in contrast to factors such as d and DT--do not have to be inferred from data, their effects will be examined by calculating percentiles of $\hat{\tau}$ for each combination of G and N:p. These percentiles should provide insight into the functioning of these two tests.

Effect of T, N:p, and G. Percentiles for $\underline{J}$ and $\underline{H}_m^*$ are displayed in Table 15. Using Bradley's liberal criterion ($.5\alpha \leq \hat{\tau} \leq 1.5\alpha$), the following patterns emerge regarding control of Type I error rates for the Johansen test and Wilcox generalization: (a) the Johansen test ($\underline{J}$) was adequate only when N:p was 20 and G was 3; and (b) the Wilcox generalization was inadequate over the range of experimental conditions considered in the experiment.

Since the performance of the Johansen test and the Wilcox generalization was so inadequate, further analysis was not warranted for either of these two tests.

Summary

It is clear that in terms of controlling Type I error rates under the heteroscedastic experimental conditions considered (a) the four Brown-Forsythe generalizations are much more effective than either the modified Wilcox test or the Johansen test, (b) the Johansen test is more effective than the modified Wilcox test, and (c) the modified Wilcox

Table 15

Percentiles of $\hat{\tau}$ for the Johansen Test (J) and Wilcox Generalization (H_m^*) for Combinations of Ratio of Total Sample Size to Number of Dependent Variables (N:p) and Number of Populations Sampled (G)

		G (N:p=10)		G (N:p=20)	
Test	Percentile	3	6	3	6
J	95th	.1700*	.7535*	.1085*	.2950*
	90th	.1260*	.6785*	.0915*	.2520*
	75th	.1030*	.5590*	.0733	.1850*
	50th	.0765*	.4548*	.0595	.1568*
	25th	.0648	.3688*	.0525	.1103*
	10th	.0580	.3245*	.0480	.0930*
	5th	.0550	.2780*	.0460	.0865*
H_m^*	95th	.2260*	.9690*	.1400*	.6795*
	90th	.1985*	.9555*	.1165*	.6170*
	75th	.1560*	.8113*	.0905*	.4435*
	50th	.1243*	.7240*	.0798*	.3263*
	25th	.0900*	.5820*	.0630	.2285*
	10th	.0760*	.5135*	.0515	.1775*
	5th	.0735	.4575*	.0490	.1535*

Note. Percentiles denoted by an asterisk fall outside the interval $[.5\alpha, 1.5\alpha]$.

test is not sufficiently effective over the set of experimental conditions considered to warrant its use.

CHAPTER 5
CONCLUSIONS

General Observations

Two hundred and fifty-six simulated conditions were investigated in a complete factorial experiment. The results obtained may be applied to experiments which have experimental conditions similar to the 256 simulated experiments conducted in this study. The generalizability of the results is limited by the range of values for the Distribution Type (DT), Number of Dependent Variables (p), Number of Populations Sampled (G), Ratio of Total Sample Size to Number of Dependent Variables (N:p), Degree of the Sample Size Ratio (NR), Form of the Sample Size Ratio (NRF), Degree of Heteroscedasticity (d), and Relationship between Sample Size and Dispersion Matrices (S). With these limitations in mind, the following conclusions may be set forth:

Conclusion 1. The estimated Type I error rates for the first modified Hotelling-Lawley test ($\underline{U}_1^*$), second modified Hotelling-Lawley test ($\underline{U}_2^*$), modified Pillai-Bartlett test ($\underline{V}^*$), and modified Wilks test ($\underline{L}^*$) were much closer to the nominal Type I error rate over the variety of conditions considered in the experiment than either the Johansen test ($\underline{J}$) or the modified Wilcox test ($\underline{H}_m^*$).

Conclusion 2. The modified Wilks test ($\underline{L}^*$) is the most effective of the Brown-Forsythe generalizations at maintaining Type I error rates. When both (a) N:p is small and (b) G is large, however, the test becomes conservative. Under conditions where both (a) N:p is small and (b) G is large, the first modified Hotelling-Lawley test is most effective at maintaining acceptable Type I error rates.

Conclusion 3. If one can tolerate a somewhat liberal test the first modified Hotelling-Lawley test ($\underline{U}_1^*$) might be used, although the shortcoming of this procedure is the test appears to become more liberal as N:p increases.

Conclusion 4. For all four of the Brown-Forsythe generalizations $\hat{\tau}$ increases as N:p increases, suggesting the procedures may not work well with very large sample sizes.

Conclusion 5. The Wilcox generalization and Johansen's test are inadequate in controlling Type I error rates over the range of experimental conditions considered.

Suggestions to Future Researchers

The generalizability of the study is limited by (a) the limited number of distributions considered, (b) the limited forms of dispersion matrices considered, and (c) the limited variation in the degree of the sample size ratios. First, the inclusion of the normal distribution and the exponential distribution gave representation to a symmetric and an extremely skewed distribution. Further research needs to be

conducted to see how these tests perform and how factors included in the study are affected when (a) data are sampled from moderately skewed distributions and (b) data are sampled from mixed distributions. It is reasonable to assume that the performance of the tests will fall somewhere between the two extremes, however, this certainly needs to be confirmed empirically. Findings in this area might be additionally strengthened by considering differing levels of kurtosis, as well. Second, only limited forms of dispersion matrices were considered. Because of the invariance properties of the test statistics, the results should be highly generalizable when data are sampled from populations with heteroscedatic dispersion matrices limited to two forms. Further research is needed to examine the influence of greater varieties of dispersion heteroscedasticity. Although the Brown-Forsythe generalizations have acceptable Type I error rates under heteroscedasticity conditions, it remains an open question whether acceptable error rates will result under homoscedastic conditions, since only heteroscedastic conditions were considered in the experiment. Since the Brown-Forsythe generalizations become slightly conservative as the degree of heteroscedasticity decreases, this suggests the tests will be even more conservative under homoscedastic conditions. These tests need to be examined empirically under homoscedastic conditions to insure this conservative trend is within acceptable tolerances. Third, since (a) the degree of the

sample size ratios (NR) is typically a factor which significantly influences Type I error rates and (b) NR was not a strong influence in this study, this suggests that greater variation in this factor should be examined to confirm its influence is indeed minimal when using the Brown-Forsythe generalizations.

It is clear that the power of the Brown-Forsythe generalizations needs to be investigated. It would be prudent to compare the power of the first modified Hotelling-Lawley test, second modified Hotelling-Lawley test, modified Wilks test, and modified Pillai-Bartlett test with that of James's (1954) second-order test, since James's test also tends to be conservative.

Finally, since the Johansen test is based upon asymptotic theory, it might be fruitful to examine the test using (a) larger sample sizes and (b) moderately skewed distributions. The test may prove to be useful if the boundary conditions where robustness no longer occurs can be more clearly defined. Using larger sample sizes may lower Type I error rates sufficiently to warrant further examination of the power of the test.

APPENDIX
ESTIMATED TYPE I ERROR RATES

Table 16, Table 17, and Table 18 include estimated Type I error rates ($\hat{\tau}$) for the First Modified Hotelling-Lawley Test ($\underline{U}_1^*$), Second Modified Hotelling-Lawley Test ($\underline{U}_2^*$), Modified Pillai-Bartlett Test ($\underline{V}^*$), and Modified Wilks Test ($\underline{L}^*$) at nominal Type I error rates (α) of .01, .05, and .10, respectively, for differing levels of Distribution Type (DT), Number of Dependent Variables (p), Number of Populations Sampled (G), Ratio of Total Sample Size to Number of Dependent Variables (p), Form of the Sample Size Ratio (NRF), Degree of the Sample Size Ratio (NR), Relationship of Sample Size to Dispersion Matrices (S), and Degree of Heteroscedasticity (d) The levels of the factors are the same as those found in Chapter 3 (see p.54) with the following modifications: (a) for DT, "E" denotes the sampled distribution was exponential whereas "N" denotes the sampled distribution was normal; (b) for S, "0" denotes the positive condition and "1" denotes the negative condition; for NR, "1" denotes the degree of the sample size ratio was 1.3 whereas "2" denotes the degree of the sample size ratio was 2; (c) for Degree of Heteroscedasticity d^2 was recorded rather then d; and (d) estimated Type I error rates are of the form 10000 times $\hat{\tau}$.

Hence, for example, the estimated Type I error rate was .0075 for the first modified Hotelling-Lawley test when $\alpha=.01$ when (a) data were sampled from an exponential distribution, (b) there were 3 dependent variables, (c) data were sampled from 3 populations, (d) the ratio of total sample size to number of dependent variables was 10, (e) data were sampled as in the first form of the sample size ratio, (f) the degree of the sample size ratio was 1.3, (g) the relationship of sample size to dispersion matrices was the positive condition, and (h) the degree of heteroscedasticity was $\sqrt{2}$.

Table 16

Estimated Type I Error Rates When α=.01

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	10	1	1	0	2	0075	0050	0045	0060	0110	0280
E	3	3	10	1	1	0	9	0325	0285	0105	0255	0270	0490
E	3	3	10	1	1	1	2	0085	0035	0030	0050	0130	0350
E	3	3	10	1	1	1	9	0095	0060	0055	0085	0560	0945
E	3	3	10	1	2	0	2	0060	0045	0030	0025	0105	0315
E	3	3	10	1	2	0	9	0245	0180	0100	0175	0145	0440
E	3	3	10	1	2	1	2	0085	0055	0045	0060	0235	0555
E	3	3	10	1	2	1	9	0110	0075	0080	0085	0810	1650
E	3	3	10	2	1	0	2	0090	0060	0045	0065	0135	0335
E	3	3	10	2	1	0	9	0130	0090	0045	0085	0460	0760
E	3	3	10	2	1	1	2	0070	0045	0025	0035	0145	0350
E	3	3	10	2	1	1	9	0405	0335	0150	0280	0575	0890
E	3	3	10	2	2	0	2	0075	0060	0025	0050	0175	0310
E	3	3	10	2	2	0	9	0120	0085	0025	0075	0350	0590
E	3	3	10	2	2	1	2	0105	0080	0055	0095	0400	0610
E	3	3	10	2	2	1	9	0370	0265	0125	0245	0870	1215
E	3	3	20	1	1	0	2	0145	0115	0090	0115	0165	0210
E	3	3	20	1	1	0	9	0240	0215	0165	0205	0110	0185
E	3	3	20	1	1	1	2	0110	0095	0090	0110	0165	0255
E	3	3	20	1	1	1	9	0125	0115	0085	0105	0350	0505

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	20	1	2	0	2	0080	0065	0055	0075	0115	0155
E	3	3	20	1	2	0	9	0295	0265	0155	0250	0140	0235
E	3	3	20	1	2	1	2	0095	0095	0075	0090	0225	0315
E	3	3	20	1	2	1	9	0140	0120	0105	0135	0555	0740
E	3	3	20	2	1	0	2	0090	0070	0075	0090	0110	0175
E	3	3	20	2	1	0	9	0115	0085	0065	0085	0290	0380
E	3	3	20	2	1	1	2	0125	0090	0080	0105	0200	0235
E	3	3	20	2	1	1	9	0400	0360	0200	0330	0325	0405
E	3	3	20	2	2	0	2	0065	0055	0035	0050	0090	0130
E	3	3	20	2	2	0	9	0140	0105	0085	0095	0195	0250
E	3	3	20	2	2	1	2	0130	0100	0060	0085	0330	0405
E	3	3	20	2	2	1	9	0430	0370	0220	0340	0460	0570
E	3	6	10	1	1	0	2	0060	0020	0025	0030	1140	3530
E	3	6	10	1	1	0	9	0260	0115	0050	0095	1760	4205
E	3	6	10	1	1	1	2	0145	0060	0035	0060	1745	4010
E	3	6	10	1	1	1	9	0085	0015	0020	0020	2780	5280
E	3	6	10	1	2	0	2	0090	0025	0015	0015	2815	6335
E	3	6	10	1	2	0	9	0170	0055	0015	0040	2575	6350
E	3	6	10	1	2	1	2	0090	0020	0005	0015	3235	6835
E	3	6	10	1	2	1	9	0060	0010	0005	0010	4895	7700
E	3	6	10	2	1	0	2	0060	0020	0005	0015	2360	5140

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	10	2	1	0	9	0100	0020	0005	0020	3105	5825
E	3	6	10	2	1	1	2	0085	0045	0005	0035	2840	5630
E	3	6	10	2	1	1	9	0200	0055	0020	0055	3975	6460
E	3	6	10	2	2	0	2	0090	0015	0015	0025	1570	3855
E	3	6	10	2	2	0	9	0105	0035	0030	0035	2040	4430
E	3	6	10	2	2	1	2	0090	0005	0025	0025	2345	4550
E	3	6	10	2	2	1	9	0275	0100	0050	0095	3810	5870
E	3	6	20	1	1	0	2	0125	0070	0035	0060	0340	1230
E	3	6	20	1	1	0	9	0335	0180	0115	0165	0555	1655
E	3	6	20	1	1	1	2	0125	0060	0045	0070	0590	1515
E	3	6	20	1	1	1	9	0135	0085	0040	0055	1230	2400
E	3	6	20	1	2	0	2	0120	0060	0035	0060	0555	3495
E	3	6	20	1	2	0	9	0265	0170	0100	0140	0585	3795
E	3	6	20	1	2	1	2	0155	0080	0070	0075	1200	3800
E	3	6	20	1	2	1	9	0125	0050	0025	0050	1990	4770
E	3	6	20	2	1	0	2	0170	0075	0040	0065	0415	0930
E	3	6	20	2	1	0	9	0175	0095	0080	0090	0740	1470
E	3	6	20	2	1	1	2	0130	0065	0065	0060	0665	1295
E	3	6	20	2	1	1	9	0345	0165	0105	0145	1035	1630
E	3	6	20	2	2	0	2	0125	0060	0055	0065	0625	1190
E	3	6	20	2	2	0	9	0165	0085	0055	0065	0635	1345

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	20	2	2	1	2	0090	0055	0035	0050	1145	1775
E	3	6	20	2	2	1	9	0380	0140	0070	0135	1770	2455
E	6	3	10	1	1	0	2	0090	0085	0060	0090	0140	0445
E	6	3	10	1	1	0	9	0355	0310	0155	0255	0290	0740
E	6	3	10	1	1	1	2	0090	0085	0050	0080	0240	0565
E	6	3	10	1	1	1	9	0085	0075	0045	0045	0560	1305
E	6	3	10	1	2	0	2	0070	0055	0050	0060	0225	0550
E	6	3	10	1	2	0	9	0305	0280	0145	0235	0195	0630
E	6	3	10	1	2	1	2	0075	0050	0060	0055	0315	0980
E	6	3	10	1	2	1	9	0130	0090	0095	0100	0815	1965
E	6	3	10	2	1	0	2	0080	0075	0070	0070	0135	0465
E	6	3	10	2	1	0	9	0125	0115	0085	0100	0395	0890
E	6	3	10	2	1	1	2	0080	0075	0055	0065	0210	0670
E	6	3	10	2	1	1	9	0320	0290	0095	0240	0475	1005
E	6	3	10	2	2	0	2	0080	0075	0070	0075	0200	0540
E	6	3	10	2	2	0	9	0135	0135	0095	0125	0270	0710
E	6	3	10	2	2	1	2	0105	0090	0105	0110	0435	1010
E	6	3	10	2	2	1	9	0235	0185	0020	0105	0885	1595
E	6	3	20	1	1	0	2	0110	0110	0100	0105	0165	0285
E	6	3	20	1	1	0	9	0255	0255	0190	0235	0145	0275
E	6	3	20	1	1	1	2	0125	0110	0075	0105	0160	0340

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	3	20	1	1	1	9	0125	0120	0105	0110	0280	0595
E	6	3	20	1	2	0	2	0085	0075	0080	0085	0110	0210
E	6	3	20	1	2	0	9	0295	0260	0215	0240	0160	0285
E	6	3	20	1	2	1	2	0085	0080	0065	0080	0230	0395
E	6	3	20	1	2	1	9	0135	0130	0120	0130	0490	0910
E	6	3	20	2	1	0	2	0125	0120	0110	0115	0105	0255
E	6	3	20	2	1	0	9	0155	0140	0100	0135	0240	0435
E	6	3	20	2	1	1	2	0105	0100	0085	0085	0165	0330
E	6	3	20	2	1	1	9	0315	0300	0180	0245	0235	0440
E	6	3	20	2	2	0	2	0105	0105	0090	0105	0155	0300
E	6	3	20	2	2	0	9	0105	0095	0070	0095	0150	0275
E	6	3	20	2	2	1	2	0110	0100	0105	0100	0255	0475
E	6	3	20	2	2	1	9	0465	0405	0190	0330	0460	0740
E	6	6	10	1	1	0	2	0160	0055	0045	0055	2380	6865
E	6	6	10	1	1	0	9	0255	0120	0015	0070	2640	7210
E	6	6	10	1	1	1	2	0105	0050	0030	0055	2995	7305
E	6	6	10	1	1	1	9	0065	0015	0005	0015	4445	8010
E	6	6	10	1	2	0	2	0100	0040	0025	0035	5255	9405
E	6	6	10	1	2	0	9	0330	0160	0025	0100	4840	9445
E	6	6	10	1	2	1	2	0100	0035	0030	0040	6650	9485
E	6	6	10	1	2	1	9	0085	0015	0010	0020	8185	9810

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	10	2	1	0	2	0080	0020	0015	0025	2410	6075
E	6	6	10	2	1	0	9	0150	0065	0015	0040	3000	6510
E	6	6	10	2	1	1	2	0145	0030	0030	0035	3435	6860
E	6	6	10	2	1	1	9	0230	0065	0015	0030	4275	7320
E	6	6	10	2	2	0	2	0070	0035	0015	0030	3100	6460
E	6	6	10	2	2	0	9	0140	0045	0020	0045	2885	6435
E	6	6	10	2	2	1	2	0120	0035	0020	0040	4565	7520
E	6	6	10	2	2	1	9	0145	0035	0010	0015	6075	8315
E	6	6	20	1	1	0	2	0120	0085	0055	0075	0445	2545
E	6	6	20	1	1	0	9	0320	0210	0110	0180	0565	2930
E	6	6	20	1	1	1	2	0075	0055	0045	0050	0620	2495
E	6	6	20	1	1	1	9	0130	0100	0070	0095	1160	3300
E	6	6	20	1	2	0	2	0190	0110	0085	0115	0715	5320
E	6	6	20	1	2	0	9	0330	0250	0100	0195	0680	5945
E	6	6	20	1	2	1	2	0110	0060	0045	0060	0865	5250
E	6	6	20	1	2	1	9	0145	0070	0040	0070	1725	6075
E	6	6	20	2	1	0	2	0155	0115	0095	0125	0500	1685
E	6	6	20	2	1	0	9	0210	0140	0080	0120	0735	2055
E	6	6	20	2	1	1	2	0135	0070	0045	0060	0575	1925
E	6	6	20	2	1	1	9	0390	0240	0070	0175	0965	2205
E	6	6	20	2	2	0	2	0130	0100	0070	0100	0780	2110

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	20	2	2	0	9	0160	0120	0095	0120	0650	2040
E	6	6	20	2	2	1	2	0145	0075	0055	0085	1270	2715
E	6	6	20	2	2	1	9	0380	0205	0045	0110	1810	3325
N	3	3	10	1	1	0	2	0090	0065	0045	0075	0130	0300
N	3	3	10	1	1	0	9	0150	0115	0065	0090	0080	0170
N	3	3	10	1	1	1	2	0105	0070	0065	0085	0130	0240
N	3	3	10	1	1	1	9	0080	0050	0030	0040	0145	0365
N	3	3	10	1	2	0	2	0115	0085	0070	0090	0165	0370
N	3	3	10	1	2	0	9	0195	0160	0065	0150	0130	0270
N	3	3	10	1	2	1	2	0100	0055	0030	0060	0190	0550
N	3	3	10	1	2	1	9	0080	0050	0020	0040	0230	0775
N	3	3	10	2	1	0	2	0120	0080	0055	0075	0135	0260
N	3	3	10	2	1	0	9	0140	0120	0080	0115	0125	0270
N	3	3	10	2	1	1	2	0090	0075	0050	0065	0125	0310
N	3	3	10	2	1	1	9	0200	0135	0050	0120	0225	0420
N	3	3	10	2	2	0	2	0115	0100	0060	0090	0110	0240
N	3	3	10	2	2	0	9	0145	0120	0075	0100	0120	0265
N	3	3	10	2	2	1	2	0090	0050	0040	0050	0260	0525
N	3	3	10	2	2	1	9	0055	0035	0010	0025	0350	0620
N	3	3	20	1	1	0	2	0100	0090	0070	0090	0090	0115
N	3	3	20	1	1	0	9	0200	0185	0110	0155	0085	0140

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	3	20	1	1	1	2	0095	0085	0085	0095	0080	0125
N	3	3	20	1	1	1	9	0115	0090	0065	0095	0100	0210
N	3	3	20	1	2	0	2	0050	0050	0040	0045	0075	0100
N	3	3	20	1	2	0	9	0210	0185	0150	0180	0115	0175
N	3	3	20	1	2	1	2	0100	0080	0050	0075	0100	0190
N	3	3	20	1	2	1	9	0105	0090	0070	0095	0115	0205
N	3	3	20	2	1	0	2	0095	0085	0075	0080	0080	0120
N	3	3	20	2	1	0	9	0140	0115	0100	0125	0120	0150
N	3	3	20	2	1	1	2	0145	0105	0090	0115	0115	0190
N	3	3	20	2	1	1	9	0125	0105	0060	0095	0095	0135
N	3	3	20	2	2	0	2	0140	0120	0110	0135	0095	0145
N	3	3	20	2	2	0	9	0165	0130	0095	0125	0095	0135
N	3	3	20	2	2	1	2	0110	0095	0070	0090	0070	0160
N	3	3	20	2	2	1	9	0175	0115	0050	0085	0145	0230
N	3	6	10	1	1	0	2	0150	0030	0015	0030	1135	3300
N	3	6	10	1	1	0	9	0250	0070	0035	0095	1135	3325
N	3	6	10	1	1	1	2	0140	0050	0015	0045	1345	3470
N	3	6	10	1	1	1	9	0165	0025	0025	0040	1595	3935
N	3	6	10	1	2	0	2	0175	0060	0010	0040	2380	5805
N	3	6	10	1	2	0	9	0255	0080	0025	0065	2165	5515
N	3	6	10	1	2	1	2	0155	0025	0035	0040	2905	6195

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	10	1	2	1	9	0080	0025	0015	0015	3835	6935
N	3	6	10	2	1	0	2	0140	0050	0020	0045	1950	4605
N	3	6	10	2	1	0	9	0170	0055	0020	0065	1895	4670
N	3	6	10	2	1	1	2	0185	0075	0040	0075	2195	4910
N	3	6	10	2	1	1	9	0160	0065	0030	0060	2795	5430
N	3	6	10	2	2	0	2	0155	0055	0035	0055	1335	3385
N	3	6	10	2	2	0	9	0205	0050	0035	0070	1260	3365
N	3	6	10	2	2	1	2	0130	0050	0020	0035	1915	4105
N	3	6	10	2	2	1	9	0190	0050	0035	0065	2710	4725
N	3	6	20	1	1	0	2	0115	0060	0045	0055	0260	0820
N	3	6	20	1	1	0	9	0290	0180	0080	0165	0175	0870
N	3	6	20	1	1	1	2	0135	0085	0075	0090	0235	0835
N	3	6	20	1	1	1	9	0145	0095	0065	0095	0315	0865
N	3	6	20	1	2	0	2	0165	0085	0035	0070	0365	2825
N	3	6	20	1	2	0	9	0355	0260	0165	0235	0370	2975
N	3	6	20	1	2	1	2	0160	0080	0030	0085	0525	2965
N	3	6	20	1	2	1	9	0170	0065	0030	0070	0785	2995
N	3	6	20	2	1	0	2	0215	0125	0080	0105	0260	0630
N	3	6	20	2	1	0	9	0220	0130	0065	0105	0220	0610
N	3	6	20	2	1	1	2	0150	0090	0085	0100	0295	0745
N	3	6	20	2	1	1	9	0250	0125	0045	0095	0355	0770

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	20	2	2	0	2	0210	0120	0095	0120	0415	0930
N	3	6	20	2	2	0	9	0255	0140	0110	0140	0345	0735
N	3	6	20	2	2	1	2	0125	0045	0025	0030	0705	1130
N	3	6	20	2	2	1	9	0280	0110	0040	0095	0815	1305
N	6	3	10	1	1	0	2	0135	0135	0100	0100	0135	0435
N	6	3	10	1	1	0	9	0195	0180	0095	0120	0145	0410
N	6	3	10	1	1	1	2	0130	0120	0100	0130	0170	0530
N	6	3	10	1	1	1	9	0115	0095	0075	0085	0160	0685
N	6	3	10	1	2	0	2	0070	0050	0065	0045	0105	0485
N	6	3	10	1	2	0	9	0130	0115	0065	0100	0120	0360
N	6	3	10	1	2	1	2	0090	0070	0070	0075	0185	0695
N	6	3	10	1	2	1	9	0055	0045	0035	0040	0205	0930
N	6	3	10	2	1	0	2	0105	0080	0045	0070	0085	0360
N	6	3	10	2	1	0	9	0180	0170	0075	0130	0125	0405
N	6	3	10	2	1	1	2	0105	0095	0095	0100	0155	0510
N	6	3	10	2	1	1	9	0115	0100	0020	0060	0225	0585
N	6	3	10	2	2	0	2	0100	0075	0070	0080	0150	0500
N	6	3	10	2	2	0	9	0190	0175	0090	0130	0145	0415
N	6	3	10	2	2	1	2	0130	0125	0085	0105	0310	0880
N	6	3	10	2	2	1	9	0060	0045	0005	0020	0370	0965
N	6	3	20	1	1	0	2	0160	0150	0150	0150	0090	0200

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	3	20	1	1	0	9	0175	0145	0120	0135	0115	0220
N	6	3	20	1	1	1	2	0095	0090	0080	0085	0100	0245
N	6	3	20	1	1	1	9	0175	0170	0115	0160	0130	0285
N	6	3	20	1	2	0	2	0115	0110	0105	0110	0085	0210
N	6	3	20	1	2	0	9	0250	0230	0195	0225	0095	0185
N	6	3	20	1	2	1	2	0120	0110	0095	0100	0110	0280
N	6	3	20	1	2	1	9	0055	0045	0040	0055	0080	0285
N	6	3	20	2	1	0	2	0125	0125	0120	0125	0115	0250
N	6	3	20	2	1	0	9	0210	0195	0145	0190	0175	0295
N	6	3	20	2	1	1	2	0080	0080	0060	0070	0085	0190
N	6	3	20	2	1	1	9	0155	0145	0055	0100	0115	0220
N	6	3	20	2	2	0	2	0135	0125	0110	0120	0140	0285
N	6	3	20	2	2	0	9	0160	0155	0140	0145	0115	0205
N	6	3	20	2	2	1	2	0085	0065	0050	0060	0130	0295
N	6	3	20	2	2	1	9	0125	0105	0050	0080	0095	0300
N	6	6	10	1	1	0	2	0195	0090	0040	0080	1865	6015
N	6	6	10	1	1	0	9	0300	0135	0015	0080	1605	5895
N	6	6	10	1	1	1	2	0150	0055	0050	0060	2115	6375
N	6	6	10	1	1	1	9	0235	0070	0010	0060	2735	6835
N	6	6	10	1	2	0	2	0185	0090	0070	0095	4590	9115
N	6	6	10	1	2	0	9	0395	0190	0050	0140	4455	9185

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	10	1	2	1	2	0135	0045	0015	0030	5475	9350
N	6	6	10	1	2	1	9	0205	0030	0015	0045	7065	9600
N	6	6	10	2	1	0	2	0135	0075	0045	0070	1740	5250
N	6	6	10	2	1	0	9	0220	0105	0025	0085	1630	5105
N	6	6	10	2	1	1	2	0190	0070	0040	0085	2365	5670
N	6	6	10	2	1	1	9	0305	0095	0000	0045	3085	6370
N	6	6	10	2	2	0	2	0190	0090	0090	0105	2530	5610
N	6	6	10	2	2	0	9	0295	0165	0070	0130	1625	4905
N	6	6	10	2	2	1	2	0180	0065	0025	0050	3745	6710
N	6	6	10	2	2	1	9	0175	0035	0000	0020	4865	7510
N	6	6	20	1	1	0	2	0185	0115	0090	0125	0220	1815
N	6	6	20	1	1	0	9	0355	0290	0125	0235	0270	1860
N	6	6	20	1	1	1	2	0175	0115	0105	0110	0240	1785
N	6	6	20	1	1	1	9	0180	0130	0075	0115	0310	1820
N	6	6	20	1	2	0	2	0155	0125	0095	0125	0390	4525
N	6	6	20	1	2	0	9	0310	0245	0100	0185	0365	4615
N	6	6	20	1	2	1	2	0195	0150	0070	0120	0420	4355
N	6	6	20	1	2	1	9	0180	0140	0075	0115	0615	4335
N	6	6	20	2	1	0	2	0140	0085	0065	0085	0135	0940
N	6	6	20	2	1	0	9	0200	0135	0085	0115	0285	1090
N	6	6	20	2	1	1	2	0210	0170	0135	0170	0405	1200

Table 16--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	20	2	1	1	9	0290	0165	0050	0115	0365	1200
N	6	6	20	2	2	0	2	0145	0130	0100	0120	0385	1335
N	6	6	20	2	2	0	9	0205	0140	0100	0130	0280	1170
N	6	6	20	2	2	1	2	0165	0130	0065	0095	0525	1685
N	6	6	20	2	2	1	9	0285	0145	0025	0090	0735	2005

Table 17

Estimated Type I Error Rates When α=.05

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	10	1	1	0	2	0395	0365	0275	0355	0530	0745
E	3	3	10	1	1	0	9	0810	0705	0555	0675	0830	1000
E	3	3	10	1	1	1	2	0375	0325	0285	0320	0715	0870
E	3	3	10	1	1	1	9	0450	0405	0355	0430	1470	1835
E	3	3	10	1	2	0	2	0345	0290	0240	0310	0635	0775
E	3	3	10	1	2	0	9	0795	0730	0550	0710	0770	0995
E	3	3	10	1	2	1	2	0350	0325	0275	0315	0980	1310
E	3	3	10	1	2	1	9	0465	0390	0310	0405	2045	2545
E	3	3	10	2	1	0	2	0390	0355	0295	0360	0640	0795
E	3	3	10	2	1	0	9	0460	0385	0330	0420	1205	1455
E	3	3	10	2	1	1	2	0430	0385	0330	0390	0675	0805
E	3	3	10	2	1	1	9	0930	0815	0590	0805	1245	1435
E	3	3	10	2	2	0	2	0375	0315	0255	0340	0665	0860
E	3	3	10	2	2	0	9	0530	0480	0345	0475	1010	1255
E	3	3	10	2	2	1	2	0500	0420	0350	0425	1165	1380
E	3	3	10	2	2	1	9	0710	0625	0450	0640	1700	1935
E	3	3	20	1	1	0	2	0565	0530	0450	0520	0560	0575
E	3	3	20	1	1	0	9	0840	0800	0660	0745	0645	0710
E	3	3	20	1	1	1	2	0465	0440	0430	0455	0680	0715
E	3	3	20	1	1	1	9	0515	0485	0415	0440	1060	1125

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	20	1	2	0	2	0400	0385	0355	0360	0505	0580
E	3	3	20	1	2	0	9	0785	0745	0690	0735	0695	0745
E	3	3	20	1	2	1	2	0440	0425	0380	0415	0880	0880
E	3	3	20	1	2	1	9	0520	0470	0465	0485	1250	1475
E	3	3	20	2	1	0	2	0475	0445	0415	0440	0575	0605
E	3	3	20	2	1	0	9	0505	0470	0460	0490	0845	0895
E	3	3	20	2	1	1	2	0495	0465	0415	0450	0620	0640
E	3	3	20	2	1	1	9	1000	0950	0795	0915	0915	0955
E	3	3	20	2	2	0	2	0400	0390	0335	0405	0615	0635
E	3	3	20	2	2	0	9	0515	0485	0445	0475	0740	0810
E	3	3	20	2	2	1	2	0590	0535	0495	0560	0890	0915
E	3	3	20	2	2	1	9	0950	0890	0665	0815	1085	1165
E	3	6	10	1	1	0	2	0290	0130	0130	0150	2780	4735
E	3	6	10	1	1	0	9	0580	0345	0240	0350	3400	5400
E	3	6	10	1	1	1	2	0360	0200	0205	0215	3420	5195
E	3	6	10	1	1	1	9	0295	0175	0120	0155	4645	6335
E	3	6	10	1	2	0	2	0310	0140	0165	0215	4555	7335
E	3	6	10	1	2	0	9	0600	0305	0175	0330	4445	7250
E	3	6	10	1	2	1	2	0310	0140	0135	0180	5310	7790
E	3	6	10	1	2	1	9	0185	0075	0095	0085	6690	8510
E	3	6	10	2	1	0	2	0330	0135	0115	0155	4175	6235

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	10	2	1	0	9	0315	0155	0115	0180	4960	6665
E	3	6	10	2	1	1	2	0295	0140	0140	0180	4795	6695
E	3	6	10	2	1	1	9	0465	0260	0155	0245	5635	7300
E	3	6	10	2	2	0	2	0340	0200	0145	0200	3345	5135
E	3	6	10	2	2	0	9	0375	0195	0150	0220	3740	5545
E	3	6	10	2	2	1	2	0325	0165	0145	0180	4085	5780
E	3	6	10	2	2	1	9	0565	0365	0215	0360	5450	6825
E	3	6	20	1	1	0	2	0505	0355	0300	0355	1135	2420
E	3	6	20	1	1	0	9	0815	0585	0430	0540	1515	2890
E	3	6	20	1	1	1	2	0520	0350	0315	0330	1670	2645
E	3	6	20	1	1	1	9	0545	0360	0280	0370	2610	3530
E	3	6	20	1	2	0	2	0465	0310	0275	0315	1650	4965
E	3	6	20	1	2	0	9	0785	0585	0425	0520	1625	5195
E	3	6	20	1	2	1	2	0540	0360	0295	0360	2390	5290
E	3	6	20	1	2	1	9	0480	0265	0210	0280	3680	6140
E	3	6	20	2	1	0	2	0515	0365	0335	0375	1355	1890
E	3	6	20	2	1	0	9	0535	0400	0340	0395	1795	2425
E	3	6	20	2	1	1	2	0540	0370	0330	0375	1710	2240
E	3	6	20	2	1	1	9	0830	0530	0405	0525	2110	2740
E	3	6	20	2	2	0	2	0500	0345	0300	0345	1630	2315
E	3	6	20	2	2	0	9	0600	0450	0335	0440	1815	2545

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	20	2	2	1	2	0495	0315	0280	0315	2380	2960
E	3	6	20	2	2	1	9	0795	0550	0315	0475	2950	3470
E	6	3	10	1	1	0	2	0445	0430	0405	0430	0700	1215
E	6	3	10	1	1	0	9	0925	0890	0670	0815	1050	1720
E	6	3	10	1	1	1	2	0375	0375	0355	0375	0760	1375
E	6	3	10	1	1	1	9	0440	0400	0370	0405	1485	2260
E	6	3	10	1	2	0	2	0455	0430	0400	0430	0720	1335
E	6	3	10	1	2	0	9	0790	0775	0595	0705	0890	1420
E	6	3	10	1	2	1	2	0405	0385	0340	0400	1235	2005
E	6	3	10	1	2	1	9	0530	0465	0430	0460	1970	3130
E	6	3	10	2	1	0	2	0410	0365	0385	0405	0680	1190
E	6	3	10	2	1	0	9	0555	0520	0445	0530	1200	1930
E	6	3	10	2	1	1	2	0390	0370	0335	0345	0865	1405
E	6	3	10	2	1	1	9	0780	0715	0455	0635	1225	1955
E	6	3	10	2	2	0	2	0345	0335	0305	0330	0780	1280
E	6	3	10	2	2	0	9	0555	0525	0370	0475	1010	1590
E	6	3	10	2	2	1	2	0480	0440	0365	0425	1260	2030
E	6	3	10	2	2	1	9	0595	0530	0235	0440	1860	2720
E	6	3	20	1	1	0	2	0550	0535	0525	0535	0640	0865
E	6	3	20	1	1	0	9	0780	0770	0620	0715	0575	0840
E	6	3	20	1	1	1	2	0510	0485	0460	0470	0695	0930

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	3	20	1	1	1	9	0525	0510	0475	0500	1025	1400
E	6	3	20	1	2	0	2	0490	0470	0460	0470	0610	0890
E	6	3	20	1	2	0	9	0805	0785	0705	0770	0615	0885
E	6	3	20	1	2	1	2	0470	0445	0440	0460	0815	1225
E	6	3	20	1	2	1	9	0570	0525	0480	0515	1280	1830
E	6	3	20	2	1	0	2	0500	0485	0495	0500	0570	0805
E	6	3	20	2	1	0	9	0610	0585	0560	0600	0885	1085
E	6	3	20	2	1	1	2	0570	0565	0535	0555	0760	1030
E	6	3	20	2	1	1	9	0855	0815	0630	0750	0805	1165
E	6	3	20	2	2	0	2	0525	0520	0500	0520	0665	0895
E	6	3	20	2	2	0	9	0560	0545	0495	0530	0725	0935
E	6	3	20	2	2	1	2	0570	0545	0505	0545	0845	1125
E	6	3	20	2	2	1	9	1045	1005	0725	0905	1220	1585
E	6	6	10	1	1	0	2	0545	0360	0285	0340	4255	7975
E	6	6	10	1	1	0	9	0745	0445	0180	0335	4665	8250
E	6	6	10	1	1	1	2	0420	0250	0195	0255	5070	8340
E	6	6	10	1	1	1	9	0355	0160	0115	0160	6405	8875
E	6	6	10	1	2	0	2	0465	0265	0210	0285	7075	9675
E	6	6	10	1	2	0	9	0885	0600	0270	0475	6785	9690
E	6	6	10	1	2	1	2	0390	0180	0190	0205	8115	9755
E	6	6	10	1	2	1	9	0315	0100	0090	0120	9105	9900

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	10	2	1	0	2	0485	0300	0230	0280	4430	7290
E	6	6	10	2	1	0	9	0560	0400	0210	0355	4880	7620
E	6	6	10	2	1	1	2	0470	0300	0230	0290	5480	7975
E	6	6	10	2	1	1	9	0570	0305	0060	0230	6050	8265
E	6	6	10	2	2	0	2	0475	0285	0195	0270	5005	7590
E	6	6	10	2	2	0	9	0495	0305	0230	0290	5090	7600
E	6	6	10	2	2	1	2	0405	0240	0190	0270	6480	8355
E	6	6	10	2	2	1	9	0430	0160	0045	0130	7535	8945
E	6	6	20	1	1	0	2	0485	0375	0340	0375	1420	4235
E	6	6	20	1	1	0	9	0860	0700	0475	0610	1685	4570
E	6	6	20	1	1	1	2	0465	0330	0315	0360	1625	4215
E	6	6	20	1	1	1	9	0625	0460	0385	0435	2520	5060
E	6	6	20	1	2	0	2	0615	0500	0455	0505	1760	7035
E	6	6	20	1	2	0	9	0925	0795	0580	0685	1930	7225
E	6	6	20	1	2	1	2	0485	0360	0335	0370	2210	6795
E	6	6	20	1	2	1	9	0625	0455	0375	0425	3370	7465
E	6	6	20	2	1	0	2	0595	0540	0485	0545	1565	3050
E	6	6	20	2	1	0	9	0700	0575	0500	0560	1965	3480
E	6	6	20	2	1	1	2	0485	0385	0305	0360	1745	3200
E	6	6	20	2	1	1	9	0885	0675	0390	0565	2065	3645
E	6	6	20	2	2	0	2	0560	0475	0380	0450	2025	3675

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	20	2	2	0	9	0600	0490	0415	0480	1885	3370
E	6	6	20	2	2	1	2	0545	0430	0380	0420	2585	4330
E	6	6	20	2	2	1	9	0835	0575	0235	0465	3135	4855
N	3	3	10	1	1	0	2	0520	0505	0410	0505	0620	0730
N	3	3	10	1	1	0	9	0555	0470	0360	0470	0435	0560
N	3	3	10	1	1	1	2	0455	0425	0370	0430	0580	0770
N	3	3	10	1	1	1	9	0420	0345	0300	0370	0660	0920
N	3	3	10	1	2	0	2	0525	0475	0370	0485	0680	0880
N	3	3	10	1	2	0	9	0570	0540	0430	0515	0580	0760
N	3	3	10	1	2	1	2	0540	0450	0335	0440	0815	1195
N	3	3	10	1	2	1	9	0285	0260	0200	0250	0935	1450
N	3	3	10	2	1	0	2	0550	0465	0420	0465	0650	0760
N	3	3	10	2	1	0	9	0505	0460	0375	0445	0555	0735
N	3	3	10	2	1	1	2	0465	0415	0320	0390	0550	0740
N	3	3	10	2	1	1	9	0515	0430	0215	0435	0740	0930
N	3	3	10	2	2	0	2	0550	0500	0380	0475	0610	0775
N	3	3	10	2	2	0	9	0505	0465	0380	0450	0535	0710
N	3	3	10	2	2	1	2	0445	0385	0330	0385	0890	1075
N	3	3	10	2	2	1	9	0290	0225	0100	0200	0975	1250
N	3	3	20	1	1	0	2	0495	0460	0445	0450	0485	0515
N	3	3	20	1	1	0	9	0610	0560	0520	0580	0515	0555

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	3	20	1	1	1	2	0425	0405	0400	0425	0460	0510
N	3	3	20	1	1	1	9	0545	0510	0455	0480	0575	0700
N	3	3	20	1	2	0	2	0435	0415	0385	0410	0450	0460
N	3	3	20	1	2	0	9	0655	0620	0565	0610	0490	0490
N	3	3	20	1	2	1	2	0460	0430	0385	0420	0495	0600
N	3	3	20	1	2	1	9	0435	0400	0345	0400	0440	0595
N	3	3	20	2	1	0	2	0530	0515	0480	0505	0465	0480
N	3	3	20	2	1	0	9	0520	0505	0500	0515	0460	0505
N	3	3	20	2	1	1	2	0610	0585	0530	0580	0605	0675
N	3	3	20	2	1	1	9	0515	0490	0360	0470	0435	0485
N	3	3	20	2	2	0	2	0545	0500	0490	0525	0525	0560
N	3	3	20	2	2	0	9	0545	0510	0490	0535	0485	0520
N	3	3	20	2	2	1	2	0485	0450	0415	0455	0535	0625
N	3	3	20	2	2	1	9	0555	0485	0320	0440	0605	0700
N	3	6	10	1	1	0	2	0575	0315	0235	0315	2505	4470
N	3	6	10	1	1	0	9	0595	0355	0260	0380	2635	4515
N	3	6	10	1	1	1	2	0590	0320	0255	0355	2825	4590
N	3	6	10	1	1	1	9	0625	0295	0205	0315	3245	5150
N	3	6	10	1	2	0	2	0485	0265	0245	0290	4095	6860
N	3	6	10	1	2	0	9	0770	0450	0225	0385	3755	6760
N	3	6	10	1	2	1	2	0570	0280	0215	0330	4660	7200

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	10	1	2	1	9	0360	0140	0180	0200	5650	7795
N	3	6	10	2	1	0	2	0485	0260	0210	0270	3745	5670
N	3	6	10	2	1	0	9	0565	0315	0275	0350	3635	5730
N	3	6	10	2	1	1	2	0475	0280	0215	0260	4020	5810
N	3	6	10	2	1	1	9	0445	0210	0110	0215	4540	6430
N	3	6	10	2	2	0	2	0595	0380	0280	0360	2955	4540
N	3	6	10	2	2	0	9	0660	0430	0305	0410	2735	4575
N	3	6	10	2	2	1	2	0530	0285	0215	0300	3615	5270
N	3	6	10	2	2	1	9	0485	0255	0150	0245	4215	5830
N	3	6	20	1	1	0	2	0620	0475	0395	0480	0865	1775
N	3	6	20	1	1	0	9	0805	0650	0405	0580	0830	1835
N	3	6	20	1	1	1	2	0545	0390	0345	0395	0915	1725
N	3	6	20	1	1	1	9	0570	0360	0310	0375	0930	1865
N	3	6	20	1	2	0	2	0625	0455	0380	0435	1180	4285
N	3	6	20	1	2	0	9	0890	0710	0545	0645	1105	4390
N	3	6	20	1	2	1	2	0655	0450	0360	0445	1360	4270
N	3	6	20	1	2	1	9	0655	0435	0350	0405	1725	4480
N	3	6	20	2	1	0	2	0625	0510	0430	0495	0925	1350
N	3	6	20	2	1	0	9	0700	0510	0430	0505	0835	1350
N	3	6	20	2	1	1	2	0715	0495	0435	0495	1045	1540
N	3	6	20	2	1	1	9	0690	0455	0295	0450	1100	1535

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	20	2	2	0	2	0685	0490	0470	0505	1345	1775
N	3	6	20	2	2	0	9	0690	0550	0475	0530	1080	1505
N	3	6	20	2	2	1	2	0490	0375	0260	0345	1570	2055
N	3	6	20	2	2	1	9	0610	0390	0225	0325	1700	2285
N	6	3	10	1	1	0	2	0555	0510	0495	0520	0645	1140
N	6	3	10	1	1	0	9	0565	0530	0370	0455	0630	1025
N	6	3	10	1	1	1	2	0560	0515	0505	0520	0725	1265
N	6	3	10	1	1	1	9	0530	0490	0435	0480	0780	1605
N	6	3	10	1	2	0	2	0490	0450	0455	0465	0680	1195
N	6	3	10	1	2	0	9	0540	0530	0400	0490	0590	1165
N	6	3	10	1	2	1	2	0460	0435	0380	0435	0830	1530
N	6	3	10	1	2	1	9	0430	0350	0315	0345	0890	1985
N	6	3	10	2	1	0	2	0495	0470	0430	0445	0595	1020
N	6	3	10	2	1	0	9	0545	0530	0415	0475	0600	1045
N	6	3	10	2	1	1	2	0510	0455	0405	0455	0710	1265
N	6	3	10	2	1	1	9	0570	0470	0190	0335	0800	1445
N	6	3	10	2	2	0	2	0505	0495	0460	0470	0750	1235
N	6	3	10	2	2	0	9	0570	0550	0445	0515	0675	1100
N	6	3	10	2	2	1	2	0550	0505	0430	0500	1115	1740
N	6	3	10	2	2	1	9	0245	0205	0075	0140	1065	1805
N	6	3	20	1	1	0	2	0515	0515	0480	0510	0485	0710

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	3	20	1	1	0	9	0640	0620	0505	0575	0525	0760
N	6	3	20	1	1	1	2	0545	0515	0485	0545	0550	0780
N	6	3	20	1	1	1	9	0465	0460	0435	0460	0585	0845
N	6	3	20	1	2	0	2	0575	0565	0530	0555	0590	0835
N	6	3	20	1	2	0	9	0795	0790	0695	0780	0620	0820
N	6	3	20	1	2	1	2	0520	0505	0490	0505	0535	0790
N	6	3	20	1	2	1	9	0460	0450	0365	0405	0545	0830
N	6	3	20	2	1	0	2	0600	0595	0575	0575	0575	0730
N	6	3	20	2	1	0	9	0650	0645	0575	0625	0540	0740
N	6	3	20	2	1	1	2	0515	0485	0420	0455	0480	0700
N	6	3	20	2	1	1	9	0650	0615	0450	0545	0520	0805
N	6	3	20	2	2	0	2	0630	0605	0625	0625	0610	0795
N	6	3	20	2	2	0	9	0605	0600	0535	0565	0590	0800
N	6	3	20	2	2	1	2	0490	0480	0460	0470	0595	0825
N	6	3	20	2	2	1	9	0490	0480	0260	0385	0595	0950
N	6	6	10	1	1	0	2	0565	0420	0310	0400	3625	7260
N	6	6	10	1	1	0	9	0735	0510	0215	0410	3440	7230
N	6	6	10	1	1	1	2	0505	0345	0270	0355	4005	7450
N	6	6	10	1	1	1	9	0675	0420	0245	0380	4725	7895
N	6	6	10	1	2	0	2	0680	0460	0355	0465	6275	9555
N	6	6	10	1	2	0	9	1005	0770	0370	0595	6110	9520

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	10	1	2	1	2	0520	0265	0210	0280	7230	9650
N	6	6	10	1	2	1	9	0485	0235	0135	0240	8140	9795
N	6	6	10	2	1	0	2	0580	0395	0290	0370	3805	6485
N	6	6	10	2	1	0	9	0680	0470	0330	0425	3490	6450
N	6	6	10	2	1	1	2	0585	0400	0365	0415	4165	6945
N	6	6	10	2	1	1	9	0675	0370	0070	0235	4920	7560
N	6	6	10	2	2	0	2	0715	0495	0385	0460	4430	6870
N	6	6	10	2	2	0	9	0785	0605	0410	0555	3375	6260
N	6	6	10	2	2	1	2	0640	0430	0300	0435	5545	7775
N	6	6	10	2	2	1	9	0405	0195	0020	0110	6455	8330
N	6	6	20	1	1	0	2	0665	0590	0485	0560	0930	3255
N	6	6	20	1	1	0	9	0910	0770	0495	0665	0985	3470
N	6	6	20	1	1	1	2	0660	0565	0500	0555	0985	3270
N	6	6	20	1	1	1	9	0615	0505	0435	0490	1165	3315
N	6	6	20	1	2	0	2	0650	0565	0470	0535	1220	6320
N	6	6	20	1	2	0	9	0850	0710	0475	0615	1150	6285
N	6	6	20	1	2	1	2	0570	0490	0455	0510	1340	6170
N	6	6	20	1	2	1	9	0580	0415	0355	0415	1590	5860
N	6	6	20	2	1	0	2	0520	0405	0350	0385	0825	2110
N	6	6	20	2	1	0	9	0730	0600	0510	0595	1035	2190
N	6	6	20	2	1	1	2	0640	0550	0505	0580	1105	2320

Table 17--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	20	2	1	1	9	0770	0570	0255	0425	1065	2420
N	6	6	20	2	2	0	2	0530	0465	0375	0445	1225	2530
N	6	6	20	2	2	0	9	0720	0625	0515	0580	1050	2285
N	6	6	20	2	2	1	2	0655	0545	0455	0525	1490	3100
N	6	6	20	2	2	1	9	0660	0530	0170	0340	1745	3350

Table 18

Estimated Type I Error Rates When α=.10

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	10	1	1	0	2	0790	0740	0655	0705	1155	1230
E	3	3	10	1	1	0	9	1330	1255	1105	1255	1545	1610
E	3	3	10	1	1	1	2	0745	0705	0650	0750	1250	1285
E	3	3	10	1	1	1	9	0850	0825	0760	0830	2275	2430
E	3	3	10	1	2	0	2	0755	0700	0600	0685	1260	1325
E	3	3	10	1	2	0	9	1205	1145	1060	1140	1395	1515
E	3	3	10	1	2	1	2	0755	0695	0680	0740	1820	1935
E	3	3	10	1	2	1	9	0775	0745	0745	0815	2870	3180
E	3	3	10	2	1	0	2	0800	0750	0715	0770	1165	1225
E	3	3	10	2	1	0	9	0910	0870	0715	0795	1930	2060
E	3	3	10	2	1	1	2	0815	0775	0770	0830	1375	1355
E	3	3	10	2	1	1	9	1310	1265	1015	1225	1855	1960
E	3	3	10	2	2	0	2	0785	0745	0685	0735	1250	1345
E	3	3	10	2	2	0	9	0990	0945	0830	0925	1640	1800
E	3	3	10	2	2	1	2	0950	0880	0740	0895	1945	1995
E	3	3	10	2	2	1	9	1115	1075	0810	1035	2450	2560
E	3	3	20	1	1	0	2	1060	1045	1060	1055	1180	1110
E	3	3	20	1	1	0	9	1375	1350	1250	1315	1175	1160
E	3	3	20	1	1	1	2	0890	0850	0860	0870	1250	1190
E	3	3	20	1	1	1	9	1060	1030	0935	1005	1610	1665

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	3	20	1	2	0	2	0835	0795	0795	0815	1045	1045
E	3	3	20	1	2	0	9	1280	1260	1160	1215	1290	1205
E	3	3	20	1	2	1	2	0925	0895	0890	0915	1410	1380
E	3	3	20	1	2	1	9	1015	0940	0920	0970	2015	2080
E	3	3	20	2	1	0	2	0955	0935	0905	0940	1090	1060
E	3	3	20	2	1	0	9	1030	0990	0945	0965	1455	1425
E	3	3	20	2	1	1	2	0935	0910	0905	0925	1195	1145
E	3	3	20	2	1	1	9	1505	1460	1340	1440	1510	1465
E	3	3	20	2	2	0	2	0890	0870	0860	0875	1140	1120
E	3	3	20	2	2	0	9	0955	0935	0895	0950	1355	1335
E	3	3	20	2	2	1	2	1065	1040	1015	1040	1410	1420
E	3	3	20	2	2	1	9	1420	1365	1215	1345	1745	1750
E	3	6	10	1	1	0	2	0620	0355	0410	0415	3875	5460
E	3	6	10	1	1	0	9	1000	0700	0515	0655	4395	6010
E	3	6	10	1	1	1	2	0675	0445	0475	0480	4470	5960
E	3	6	10	1	1	1	9	0600	0340	0315	0385	5640	6885
E	3	6	10	1	2	0	2	0650	0440	0430	0460	5685	7920
E	3	6	10	1	2	0	9	1010	0705	0530	0735	5490	7830
E	3	6	10	1	2	1	2	0595	0345	0360	0415	6380	8210
E	3	6	10	1	2	1	9	0420	0220	0295	0300	7610	8815
E	3	6	10	2	1	0	2	0545	0355	0340	0370	5285	6835

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	10	2	1	0	9	0620	0375	0340	0380	5960	7215
E	3	6	10	2	1	1	2	0535	0365	0325	0380	5855	7315
E	3	6	10	2	1	1	9	0705	0530	0310	0490	6585	7725
E	3	6	10	2	2	0	2	0590	0450	0410	0470	4440	5905
E	3	6	10	2	2	0	9	0690	0420	0415	0470	4860	6225
E	3	6	10	2	2	1	2	0565	0395	0370	0455	5165	6320
E	3	6	10	2	2	1	9	0770	0630	0515	0590	6365	7300
E	3	6	20	1	1	0	2	0910	0730	0650	0745	1955	3115
E	3	6	20	1	1	0	9	1330	1130	0890	1070	2400	3650
E	3	6	20	1	1	1	2	1025	0775	0735	0800	2455	3345
E	3	6	20	1	1	1	9	0950	0755	0690	0775	3495	4265
E	3	6	20	1	2	0	2	0915	0740	0655	0735	2615	5975
E	3	6	20	1	2	0	9	1325	1100	0905	1060	2500	6005
E	3	6	20	1	2	1	2	0905	0710	0650	0700	3420	6100
E	3	6	20	1	2	1	9	0775	0585	0520	0620	4675	6810
E	3	6	20	2	1	0	2	0970	0760	0695	0785	2150	2555
E	3	6	20	2	1	0	9	0960	0785	0705	0770	2670	3155
E	3	6	20	2	1	1	2	1010	0820	0715	0805	2615	3000
E	3	6	20	2	1	1	9	1230	0955	0745	0880	3065	3415
E	3	6	20	2	2	0	2	1015	0820	0705	0790	2685	3050
E	3	6	20	2	2	0	9	1040	0865	0730	0845	2750	3255

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	3	6	20	2	2	1	2	0960	0715	0690	0780	3380	3735
E	3	6	20	2	2	1	9	1135	0895	0680	0860	3840	4180
E	6	3	10	1	1	0	2	0970	0910	0850	0915	1330	1920
E	6	3	10	1	1	0	9	1445	1400	1095	1275	1775	2280
E	6	3	10	1	1	1	2	0865	0835	0785	0820	1425	2175
E	6	3	10	1	1	1	9	0880	0830	0755	0805	2220	3090
E	6	3	10	1	2	0	2	1005	0965	0915	0940	1370	2035
E	6	3	10	1	2	0	9	1290	1270	1100	1240	1510	2010
E	6	3	10	1	2	1	2	0905	0855	0870	0910	1990	2830
E	6	3	10	1	2	1	9	0925	0860	0810	0880	2890	4025
E	6	3	10	2	1	0	2	0890	0860	0780	0850	1295	1920
E	6	3	10	2	1	0	9	1070	1050	0885	1000	2015	2555
E	6	3	10	2	1	1	2	0760	0715	0745	0800	1455	2085
E	6	3	10	2	1	1	9	1270	1200	0890	1075	1955	2635
E	6	3	10	2	2	0	2	0750	0725	0645	0715	1410	2050
E	6	3	10	2	2	0	9	1035	1000	0895	0995	1690	2375
E	6	3	10	2	2	1	2	0850	0820	0870	0850	2060	2770
E	6	3	10	2	2	1	9	1050	0965	0490	0840	2650	3520
E	6	3	20	1	1	0	2	1080	1045	1010	1045	1240	1460
E	6	3	20	1	1	0	9	1295	1280	1155	1240	1220	1480
E	6	3	20	1	1	1	2	0925	0910	0920	0915	1190	1495

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	3	20	1	1	1	9	1005	0950	0900	0945	1690	2050
E	6	3	20	1	2	0	2	1000	0985	0955	0975	1200	1530
E	6	3	20	1	2	0	9	1370	1340	1260	1300	1185	1440
E	6	3	20	1	2	1	2	1025	1005	0940	1000	1510	1910
E	6	3	20	1	2	1	9	1130	1095	1030	1095	2050	2585
E	6	3	20	2	1	0	2	0970	0960	0930	0960	1085	1330
E	6	3	20	2	1	0	9	1080	1065	0995	1050	1440	1685
E	6	3	20	2	1	1	2	1190	1170	1135	1185	1395	1545
E	6	3	20	2	1	1	9	1380	1350	1135	1220	1420	1735
E	6	3	20	2	2	0	2	1020	1000	1000	1005	1325	1600
E	6	3	20	2	2	0	9	1010	1000	0955	0975	1330	1625
E	6	3	20	2	2	1	2	1100	1075	1055	1090	1460	1860
E	6	3	20	2	2	1	9	1570	1515	1250	1450	1860	2320
E	6	6	10	1	1	0	2	0955	0740	0625	0760	5525	8460
E	6	6	10	1	1	0	9	1195	0880	0445	0685	5980	8700
E	6	6	10	1	1	1	2	0795	0585	0485	0570	6345	8815
E	6	6	10	1	1	1	9	0680	0425	0320	0425	7370	9155
E	6	6	10	1	2	0	2	0865	0615	0565	0660	7895	9790
E	6	6	10	1	2	0	9	1310	1120	0630	0920	7710	9775
E	6	6	10	1	2	1	2	0710	0485	0460	0545	8740	9845
E	6	6	10	1	2	1	9	0615	0310	0210	0310	9435	9950

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	10	2	1	0	2	0860	0670	0545	0685	5640	7845
E	6	6	10	2	1	0	9	0965	0670	0555	0695	6030	8155
E	6	6	10	2	1	1	2	0745	0575	0500	0550	6490	8455
E	6	6	10	2	1	1	9	0930	0600	0280	0455	7015	8630
E	6	6	10	2	2	0	2	0895	0680	0560	0670	6295	8130
E	6	6	10	2	2	0	9	0980	0710	0530	0690	6115	8165
E	6	6	10	2	2	1	2	0720	0505	0420	0515	7380	8750
E	6	6	10	2	2	1	9	0640	0355	0145	0285	8185	9245
E	6	6	20	1	1	0	2	1025	0875	0805	0900	2340	5260
E	6	6	20	1	1	0	9	1375	1205	0835	1060	2670	5510
E	6	6	20	1	1	1	2	0885	0790	0765	0780	2485	5155
E	6	6	20	1	1	1	9	1085	0950	0775	0915	3475	5970
E	6	6	20	1	2	0	2	1105	0975	0925	0970	2620	7905
E	6	6	20	1	2	0	9	1430	1315	1060	1215	2935	8000
E	6	6	20	1	2	1	2	0910	0805	0745	0785	3245	7665
E	6	6	20	1	2	1	9	1055	0885	0750	0870	4360	8135
E	6	6	20	2	1	0	2	1070	0970	0920	0965	2405	4050
E	6	6	20	2	1	0	9	1100	1000	0880	0995	2850	4365
E	6	6	20	2	1	1	2	1010	0880	0805	0865	2620	4165
E	6	6	20	2	1	1	9	1360	1115	0715	0975	3025	4670
E	6	6	20	2	2	0	2	1110	0995	0915	0965	2975	4570

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
E	6	6	20	2	2	0	9	1095	0970	0865	0935	2745	4395
E	6	6	20	2	2	1	2	0990	0855	0765	0860	3640	5245
E	6	6	20	2	2	1	9	1255	0955	0545	0815	4125	5620
N	3	3	10	1	1	0	2	0970	0935	0885	0920	1095	1125
N	3	3	10	1	1	0	9	0940	0890	0785	0865	0945	0975
N	3	3	10	1	1	1	2	0900	0865	0805	0890	1205	1220
N	3	3	10	1	1	1	9	0835	0795	0685	0785	1295	1485
N	3	3	10	1	2	0	2	0960	0920	0820	0915	1240	1330
N	3	3	10	1	2	0	9	0990	0940	0845	0915	1155	1210
N	3	3	10	1	2	1	2	0925	0865	0765	0875	1535	1710
N	3	3	10	1	2	1	9	0665	0585	0580	0655	1575	2015
N	3	3	10	2	1	0	2	1030	0975	0910	0980	1105	1180
N	3	3	10	2	1	0	9	0990	0905	0830	0945	1095	1155
N	3	3	10	2	1	1	2	0905	0820	0790	0855	1205	1280
N	3	3	10	2	1	1	9	0910	0865	0630	0785	1330	1415
N	3	3	10	2	2	0	2	0990	0945	0915	0970	1230	1290
N	3	3	10	2	2	0	9	0955	0905	0815	0870	1065	1155
N	3	3	10	2	2	1	2	0860	0795	0705	0835	1550	1645
N	3	3	10	2	2	1	9	0570	0525	0295	0525	1605	1870
N	3	3	20	1	1	0	2	1005	0970	0965	0980	1015	0955
N	3	3	20	1	1	0	9	1120	1100	1015	1065	1035	1010

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	3	20	1	1	1	2	1010	0985	0940	0980	1020	1010
N	3	3	20	1	1	1	9	0995	0980	0930	0960	1045	1110
N	3	3	20	1	2	0	2	0975	0970	0940	0975	1000	0965
N	3	3	20	1	2	0	9	1165	1150	1065	1140	0945	0910
N	3	3	20	1	2	1	2	0885	0855	0800	0835	1015	1025
N	3	3	20	1	2	1	9	0880	0840	0795	0855	0920	1035
N	3	3	20	2	1	0	2	1020	0990	0935	0955	0990	0945
N	3	3	20	2	1	0	9	1085	1055	0990	1030	1020	1005
N	3	3	20	2	1	1	2	1185	1160	1125	1155	1165	1120
N	3	3	20	2	1	1	9	0980	0920	0720	0875	0865	0850
N	3	3	20	2	2	0	2	1095	1085	1035	1085	1110	1045
N	3	3	20	2	2	0	9	1015	0985	0975	1000	1035	0945
N	3	3	20	2	2	1	2	0885	0855	0855	0870	1090	1120
N	3	3	20	2	2	1	9	0975	0940	0785	0870	1120	1150
N	3	6	10	1	1	0	2	1120	0775	0570	0755	3630	5155
N	3	6	10	1	1	0	9	0945	0670	0510	0650	3645	5190
N	3	6	10	1	1	1	2	1030	0745	0675	0785	3845	5290
N	3	6	10	1	1	1	9	0990	0700	0585	0715	4310	5875
N	3	6	10	1	2	0	2	0960	0620	0645	0695	5110	7435
N	3	6	10	1	2	0	9	1200	0895	0630	0865	4745	7390
N	3	6	10	1	2	1	2	0940	0640	0625	0750	5640	7720

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	10	1	2	1	9	0630	0365	0405	0495	6655	8155
N	3	6	10	2	1	0	2	0895	0595	0500	0610	4790	6265
N	3	6	10	2	1	0	9	1020	0665	0635	0725	4820	6405
N	3	6	10	2	1	1	2	0840	0560	0505	0610	5100	6355
N	3	6	10	2	1	1	9	0730	0450	0345	0445	5610	7015
N	3	6	10	2	2	0	2	0935	0715	0575	0755	3970	5210
N	3	6	10	2	2	0	9	1105	0790	0605	0790	3855	5270
N	3	6	10	2	2	1	2	0910	0645	0650	0680	4620	5915
N	3	6	10	2	2	1	9	0715	0470	0365	0465	5245	6390
N	3	6	20	1	1	0	2	1125	0970	0875	0960	1530	2545
N	3	6	20	1	1	0	9	1270	1050	0860	0985	1525	2445
N	3	6	20	1	1	1	2	1085	0860	0765	0850	1585	2470
N	3	6	20	1	1	1	9	1050	0840	0710	0860	1785	2665
N	3	6	20	1	2	0	2	1090	0905	0860	0895	1965	5160
N	3	6	20	1	2	0	9	1405	1235	1060	1160	1810	5305
N	3	6	20	1	2	1	2	1150	0925	0815	0950	2245	5080
N	3	6	20	1	2	1	9	1040	0810	0700	0850	2600	5210
N	3	6	20	2	1	0	2	1155	0930	0920	0960	1630	1930
N	3	6	20	2	1	0	9	1170	0980	0910	0985	1625	1920
N	3	6	20	2	1	1	2	1235	1050	0930	1025	1830	2145
N	3	6	20	2	1	1	9	1105	0850	0630	0805	1750	2090

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	3	6	20	2	2	0	2	1160	0985	0950	1010	2040	2355
N	3	6	20	2	2	0	9	1180	1030	0910	0970	1800	2155
N	3	6	20	2	2	1	2	1025	0770	0725	0810	2355	2705
N	3	6	20	2	2	1	9	0965	0675	0445	0645	2495	2925
N	6	3	10	1	1	0	2	1160	1130	1025	1085	1285	1880
N	6	3	10	1	1	0	9	0950	0920	0720	0850	1140	1635
N	6	3	10	1	1	1	2	1065	1005	0985	1000	1330	1960
N	6	3	10	1	1	1	9	1085	1060	0925	1040	1510	2390
N	6	3	10	1	2	0	2	0955	0920	0880	0910	1250	1820
N	6	3	10	1	2	0	9	1055	1010	0855	0955	1225	1740
N	6	3	10	1	2	1	2	0905	0880	0820	0880	1455	2220
N	6	3	10	1	2	1	9	0830	0775	0700	0795	1670	2865
N	6	3	10	2	1	0	2	0945	0930	0870	0905	1090	1605
N	6	3	10	2	1	0	9	0950	0925	0755	0865	1125	1855
N	6	3	10	2	1	1	2	0970	0935	0855	0925	1360	1945
N	6	3	10	2	1	1	9	0925	0880	0530	0805	1430	2145
N	6	3	10	2	2	0	2	1025	0995	0960	0995	1320	1895
N	6	3	10	2	2	0	9	1195	1135	0950	1080	1220	1740
N	6	3	10	2	2	1	2	1060	1015	0930	0995	1775	2450
N	6	3	10	2	2	1	9	0550	0515	0210	0355	1720	2445
N	6	3	20	1	1	0	2	1030	1020	1005	1005	1050	1260

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	3	20	1	1	0	9	1165	1155	0990	1120	1045	1285
N	6	3	20	1	1	1	2	1055	1025	0990	1020	1115	1350
N	6	3	20	1	1	1	9	0955	0920	0890	0910	1090	1440
N	6	3	20	1	2	0	2	1155	1130	1115	1135	1220	1460
N	6	3	20	1	2	0	9	1295	1295	1185	1240	1125	1340
N	6	3	20	1	2	1	2	0955	0930	0925	0940	1025	1435
N	6	3	20	1	2	1	9	0925	0915	0855	0900	0960	1515
N	6	3	20	2	1	0	2	1025	1015	1000	1020	1040	1295
N	6	3	20	2	1	0	9	1085	1070	1035	1055	0965	1215
N	6	3	20	2	1	1	2	0965	0940	0905	0925	1020	1280
N	6	3	20	2	1	1	9	1060	1020	0870	0975	1065	1430
N	6	3	20	2	2	0	2	1175	1165	1125	1150	1105	1370
N	6	3	20	2	2	0	9	1155	1145	1055	1105	1105	1375
N	6	3	20	2	2	1	2	0935	0920	0880	0915	1035	1350
N	6	3	20	2	2	1	9	0950	0905	0620	0795	1130	1580
N	6	6	10	1	1	0	2	0985	0835	0690	0795	4775	7870
N	6	6	10	1	1	0	9	1165	0895	0460	0700	4575	7910
N	6	6	10	1	1	1	2	0985	0755	0645	0755	5210	7975
N	6	6	10	1	1	1	9	1065	0780	0620	0785	5860	8390
N	6	6	10	1	2	0	2	1080	0880	0735	0850	7175	9695
N	6	6	10	1	2	0	9	1610	1260	0765	1110	7055	9695

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	10	1	2	1	2	0945	0645	0525	0650	8050	9745
N	6	6	10	1	2	1	9	0835	0495	0395	0530	8835	9860
N	6	6	10	2	1	0	2	1135	0905	0730	0900	4880	7275
N	6	6	10	2	1	0	9	1145	0890	0685	0885	4640	7230
N	6	6	10	2	1	1	2	1035	0790	0730	0800	5295	7595
N	6	6	10	2	1	1	9	0990	0675	0210	0540	5995	8085
N	6	6	10	2	2	0	2	1230	1040	0910	1045	5520	7500
N	6	6	10	2	2	0	9	1200	1000	0825	0955	4545	6940
N	6	6	10	2	2	1	2	1090	0835	0710	0815	6520	8295
N	6	6	10	2	2	1	9	0660	0355	0110	0260	7310	8675
N	6	6	20	1	1	0	2	1100	1030	0985	1030	1585	4330
N	6	6	20	1	1	0	9	1400	1275	0945	1125	1700	4585
N	6	6	20	1	1	1	2	1185	1090	1000	1090	1755	4265
N	6	6	20	1	1	1	9	1070	0930	0810	0890	1785	4320
N	6	6	20	1	2	0	2	1145	1045	0990	1040	1900	7235
N	6	6	20	1	2	0	9	1370	1290	0960	1140	1805	7150
N	6	6	20	1	2	1	2	1045	0915	0860	0925	2140	7175
N	6	6	20	1	2	1	9	1095	0900	0740	0870	2415	6830
N	6	6	20	2	1	0	2	1000	0925	0845	0900	1625	2975
N	6	6	20	2	1	0	9	1270	1165	1010	1130	1765	3095
N	6	6	20	2	1	1	2	1090	1000	0930	1020	1845	3170

Table 18--continued.

DT	p	G	N:p	NRF	NR	S	d^2	U_1^*	U_2^*	V^*	L^*	J	H_m^*
N	6	6	20	2	1	1	9	1275	1030	0610	0845	1845	3320
N	6	6	20	2	2	0	2	1030	0920	0820	0890	1990	3370
N	6	6	20	2	2	0	9	1350	1225	1045	1145	1815	3210
N	6	6	20	2	2	1	2	1115	0945	0880	0945	2400	4050
N	6	6	20	2	2	1	9	1070	0795	0385	0685	2595	4215

REFERENCES

Algina, J., & Oshima, T.C. (1990). Robustness of the independent sample Hotelling's T^2 to variance-covariance heteroscedasticity when sample sizes are unequal or in small ratios. Psychological Bulletin, 108, 308-313.

Algina, J., Oshima, T.C., & Tang, K.L. (1991). Robustness of Yao's, James', and Johansen's tests under variance-covariance heteroscedasticity and non-normality. Journal of Educational Statistics, 16(2), 125-139.

Algina, J., & Tang, K.L. (1988). Type I error rates for Yao's and James' tests of equality of mean vectors under variance-covariance heteroscedasticity. Journal of Educational Statistics, 13, 281-290.

Anderson, T.W. (1958). An introduction to multivariate statistical analysis. New York: Wiley.

Aspin, A.A. (1948). An examination and further development of formula arising in the problem of comparing two mean values. Biometrika, 35, 88-96.

Aspin, A.A. (1949). Tables for use in comparisons whose accuracy involves two variances, separately estimated (with appendix by B.L. Welch). Biometrika, 36, 290-296.

Bartlett, M.S. (1939). A note on tests of significance in multivariate analysis. Proceedings of the Cambridge Philosophical Society, 35, 180-185.

Bennet, B.M. (1951). Note on a solution of the generalized Behrens-Fisher problem. Annals of the Institute of Statistical Mathematics, 25, 290-302.

Boneau, C.A. (1960). The effects of violations of assumptions underlying the t test. Psychological Bulletin, 57(1), 49-64.

Box, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. Annals of Mathematical Statistics, 25, 290-302.

Bradley, J.V. (1978). Robustness? British Journal of Mathematical and Statistical Psychology, 31, 144-152.

Bray, J.H., & Maxwell, S.E. (1989). Multivariate analysis of variance (2nd ed.). Newbury Park, CA : Sage.

Brown, M.B., & Forsythe, A.B. (1974). The small sample behavior of some statistics which test the equality of several means. Technometrics, 16(1), 129-132.

Clinch, J.J., & Keselman, H.J. (1982). Parametric alternatives to the analysis of variance. Journal of Educational Statistics, 7, 207-214.

Cochran, W.G. (1947). Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics, 2, 22-38.

Dijkstra, J.B., & Werter, S.P.J. (1981). Testing the equality for several means when the population variances are unequal. Communication in Statistics-Simulation and Computation, B10, 557-569.

Everitt, B.S. (1979). A Monte Carlo investigation of the robustness of Hotelling's one- and two-sample T^2 tests. Journal of the American Statistical Association, 74, 48-51.

Fisher, R.A. (1935). The fiducial argument in statistical inference. Annals of Eugenics, 6, 391-398.

Fisher, R.A. (1939). The comparison of samples with possible unequal variances. Annals of Eugenics, 9, 174-180.

Glass, G.V., Peckham, P.D., & Sanders, J.R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance. Review of Educational Research, 42(3), 237-288.

Hakstian, A.R., Roed, J.C., & Lind, J.C. (1979). Two-sample T^2 procedure and the assumption of homogeneous covariance matrices. Psychological Bulletin, 86, 1255-1263.

Holloway, L.N., & Dunn, O.J. (1967). The robustness of Hotelling's T^2. Journal of the American Statistical Association, 62, 124-136.

Hopkins, J.W., & Clay, P.P.F. (1963). Some empirical distributions of bivariate T^2 and homoscedasticity criterion M under unequal variance and leptokurtosis. Journal of the American Statistical Association, 58, 1048-1053.

Horsnell, G. (1953). The effect of unequal group variances on the F-test for the homogeneity of group means. Biometrika, 40, 128-136.

Hotelling, H. (1931). The generalization of Student's ratio. Annals of Mathematical Statistics, 2, 360-378.

Hotelling, H. (1951). A generalized T test and measure of multivariate dispersion. In Jerzy Neyman (Ed.), Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability (pp. 23-41). Berkley : University of California Press.

Hsu, P.L. (1938). Contributions to the theory of 'Student's' t-test as applied to the problem of two samples. Statistical Research Memoirs, 2, 1-24.

Hughes, D.T., & Saw, J.G. (1972). Approximating the percentage points of Hotelling's generalized T_o^2 statistic. Biometrika, 59, 224-226.

Ito, K. (1969). On the effect of heteroscedasticity and non-normality upon some multivariate test procedures. In P.R. Krishnaiah (Ed.), Multivariate Analysis-II, (pp. 87-120). New York : Academic Press.

Ito, K., & Schull, W.J. (1964). On the robustness of the T^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. Biometrika, 38, 324-329.

James, G.S. (1951). The comparison of several groups of observations when the ratios of population variances are unknown. Biometrika, 38, 324-329.

James, G.S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. Biometrika, 41, 19-43.

Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. Biometrika, 67, 85-92.

Kendall, M.G., & Stuart, A. (1969). The advanced theory of statistics. Vol. 1. New York: Hafner.

Kohr, R.L., & Games, P.A. (1974). Robustness of the analysis of variance, the Welch procedure and a Box procedure to heterogeneous variables. Journal of Experimental Education, 43, 61-69.

Korin, B.P. (1972). Some comments of the homoscedasticity criterion M and the multivariate analysis of variance tests T, W, and R. Biometrika, 59, 215-216.

Lawley, D.N. (1938). A generalization of Fisher's z-test. Biometrika, 30, 180-187.

Levy, K.J. (1978). An empirical comparison of the ANOVA F-test with alternatives which are more robust against heterogeneity of variance. Journal of Statistical Computing and Simulation, 8, 49-57.

Lin, W. (1992). Robustness of two multivariate tests to variance-covariance heteroscedasticity and nonnormality when total-sample-size-to-variable ratio is small. Dissertation Abstracts International, 52(8), 2899A.

Lindquist, E.F. (1953). Design and analysis of experiments in education and psychology. Boston: Houghton Mifflin.

McKeon, J.J. (1974). F approximations to the distributions of Hotelling's T_0^2. Biometrika, 61, 381-383.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. Psychological Bulletin, 105(1), 156-166.

Nath, R., & Duran, B.S. (1983). A robust test in the multivariate two-sample location problem. American Journal of Mathematical and Management Sciences, 3, 225-249.

Nel, D.G., & van der Merwe, C.A. (1986). A solution to the multivariate Behrens-Fisher problem. Communications in Statistics - Theory and Methodology, 15(2), 3719-3735.

Novick, M.R., & Jackson, P.H. (1974). Statistical methods for educational and psychological research. New York: McGraw- Hill.

Olson, C.L. (1974). Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 69, 894-908.

Olson, C.L. (1976). On choosing a test statistic in multivariate analysis of variance. Psychological Bulletin, 83(4), 579-586.

Olson, C.L. (1979). Practical considerations in choosing a MANOVA test statistic: A rejoinder to Stevens. Psychological Bulletin, 86(6), 1350-1352.

Oshima, T.C., & Algina, J. (in press). Type I error rates for James's second-order test and Wilcox's $\underline{H}_m$ test under heteroscedasticity and non-normality. British Journal of Mathematical and Statistical Psychology.

Pearson, E.S. (1929). The distribution of frequency constants in small samples from non-normal symmetrical and skewed populations. Biometrika, 21, 259-286.

Pearson, E.S. (1931). The analysis of variance in cases of non-normal variation. Biometrika, 23, 114-133.

Pearson, E.S., & Please, N.W. (1975). Relation between the shape of population distributions and the robustness of four simple test statistics. Biometrika, 62, 223-241.

Pillai, K.C.S. (1955). Some new test criteria in multivariate analysis. Annals of Mathematical Statistics, 26, 117-121.

Pillai, K.C.S., & Sampson, P. (1959). On Hotelling's generalization of T^2. Biometrika, 46, 160-165.

Pillai, K.C.S., & Sudjana. (1975). Exact robustness studies of tests of two multivariate hypotheses based on four criteria and their distribution problems under violations. The Annals of Statistics, 3, 617-636.

Pratt, J.W. (1964). Robustness of some procedures for the two-sample location problem. Journal of the American Statistical Association, 59, 665-680.

Ramsey, P.H. (1980). Exact Type I error rates for robustness of Student's t test with unequal variances. Journal of Educational Statistics, 5, 337-349.

Rao, C.R. (1952). Advanced statistical methods in biometric research. New York: Wiley.

Ratcliffe, J.F. (1968). The effect on the t distribution of non-normality in the sampled population. Applied Statistics, 17(1), 42-48.

Rider, P.R. (1929). On the distribution of the ratio of mean to standard deviation in small samples from non-normal populations. Biometrika, 21, 124-143.

Rogan, J.C., & Keselman, H.J. (1977). Is then ANOVA F-test robust to variance heterogeneity when sample sizes are equal?: An investigation via coefficient of variation. American Educational Research Journal, 14, 493-498.

Roy, S.N. (1945). The individual sampling distribution of the maximum, the minimum, and any intermediate of the p-statistics on the null-hypothesis. Sankhya, 7, Part 2, 133-158.

SAS Institute Inc. (1985). SAS user's guide, 1985 edition. Raleigh, NC: Author.

Satterthwaite, F.E. (1941). Synthesis of variance. Psychometrika, 6, 309-316.

Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. Biometrics, 2(6), 110-114.

Scheffe, H. (1943). On solutions of the Behrens-Fisher problem based on the t-distribution. Annals of Mathematical Statistics, 14, 35-44.

Scheffe, H. (1959). Analysis of variance. New York: John Wiley and Sons.

Scheffe, H. (1970). Practical solutions to the Behrens-Fisher problem. Journal of the American Statistical Association, 65, 1051-1058.

Stevens, J. (1979). Comment on Olson: Choosing a test statistic in multivariate analysis of variance. Psychological Bulletin, 86(2), 355-360.

Stevens, J. (1986). Applied multivariate statistics for the social sciences. Hillsdale, NJ : Lawrence Erlbaum Associates.

Tang, K.L. (1990). Robustness of four multivariate tests under variance-covariance heteroscedasticity. Dissertation Abstracts International, 51(5), 2444B.

Tomarken, A., & Serlin, R. (1986). Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. Psychological Bulletin, 99, 90-99.

Wang, Y.Y. (1971). Probabilities of the type I errors of the Welch tests for the Behrens-Fisher problem. Journal of the American Statistical Association, 66, 605-608.

Welch, B.L. (1938). The significance of the difference between two means when the population variances are unequal. Biometrika, 29, 350-362.

Welch, B.L. (1947). The generalization of 'Students' problem when several different population variances are involved. Biometrika, 34, 23-35.

Welch, B.L. (1951). On the comparison of several mean values: an alternative approach. Biometrika, 38, 330-336.

Wilcox, R.R. (1988). A new alternative to the ANOVA F and new results on James' second-order method. British Journal of Mathematical and Statistical Psychology, 41, 109-117.

Wilcox, R.R. (1989). Adjusting for unequal variances when comparing means in one-way and two-way fixed effects ANOVA models. Journal of Educational Statistics, 14(3), 269-278.

Wilcox, R.R. (1990). Comparing the means of two independent groups. Biometric Journal, 32, 771-780.

Wilcox, R.R., Charlin, & Thompson, (1986). New Monte Carlo results on the robustness of the ANOVA F, W, and F^* statistics. Communications in Statistics-Simulation and Computation, 15, 933-944.

Wilks, S.S. (1932) Certain generalizations in the analysis of variance. Biometrika, 24, 471-494.

Yao, Y. (1965). An approximate degrees of freedom solution to the multivariate Behrens-Fisher problem. Biometrika, 52, 139-147.

Young, R.K., & Veldman, D.J. (1965). Introductory statistics for the behavioral sciences. New York: Holt, Rinehart & Winston.

Yuen, K.K. (1974). The two-sample trimmed t for unequal population variances. Biometrika, 61, 165-176.

BIOGRAPHICAL SKETCH

William Thomas Coombs was born September 30, 1954. He received two bachelor's degrees with majors in history (1976) and psychology (1979), both from the University of Tennessee. He next received three master's degrees with majors in human relations (1981), mathematics (1987), and statistics (1989), from Shippensburg State College, Bowling Green State University, and the University of Florida, respectively.

In the fall of 1989, he began studying for the Ph.D. degree in the Foundations of Education Department at the University of Florida, majoring in research and evaluation methodology. He will graduate with the Ph.D. degree in August, 1992, and begin his career as an Assistant Professor in Applied Behavioral Studies at Oklahoma State University.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James Algina

James J. Algina, Chair
Professor of Foundations of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Linda Crocker

Linda M. Crocker
Professor of Foundations of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

M. David Miller

M. David Miller
Associate Professor of Foundations of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald H. Randles

Ronald H. Randles
Professor of Statistics

This dissertation was submitted to the Graduate Faculty of the College of Education and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1992

James Ole...

Chairperson, Foundations of Education

David C. Smith

Dean, College of Education

Dean, Graduate School